ENDOZEPINE-LIKE PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF USING THE SAME

RELATED APPLICATIONS

This application claims priority to U.S.S.N. 60/253,834, filed November 29, 2000; U.S.S.N. 60/264,180, filed January 25, 2001; and U.S.S.N. 60/313,656, filed August 20, 2001, each of which is incorporated by reference in its entirety.

FIELD OF THE INVENTION

The invention relates to endozepine-like polynucleotides and the polypeptides encoded by such polynucleotides, as well as vectors, host cells, antibodies and recombinant methods for producing the polypeptides and polynucleotides, as well as methods for using the same.

BACKGROUND OF THE INVENTION

The invention generally relates to nucleic acids and polypeptides encoded therefrom. More specifically, the invention relates to nucleic acids encoding cytoplasmic, nuclear, membrane bound, and secreted polypeptides, as well as vectors, host cells, antibodies, and recombinant methods for producing these nucleic acids and polypeptides.

SUMMARY OF THE INVENTION

The invention is based in part upon the discovery of nucleic acid sequences encoding novel polypeptides. The novel nucleic acids and polypeptides are referred to herein as NOVX, NOV1 or NOV2 nucleic acids and polypeptides. These nucleic acids and polypeptides, as well as derivatives, homologs, analogs and fragments thereof, will hereinafter be collectively designated as "NOVX" nucleic acid or polypeptide sequences.

In one aspect, the invention provides an isolated NOVX nucleic acid molecule encoding a NOVX polypeptide that includes a nucleic acid sequence that has identity to the nucleic acids disclosed in SEQ ID NO: 1, 3, 4, 6, 8, or 9. In some embodiments, the NOVX nucleic acid molecule will hybridize under stringent conditions to a nucleic acid sequence

complementary to a nucleic acid molecule that includes a protein-coding sequence of a NOVX nucleic acid sequence. The invention also includes an isolated nucleic acid that encodes a NOVX polypeptide, or a fragment, homolog, analog or derivative thereof. For example, the nucleic acid can encode a polypeptide at least 80% identical to a polypeptide comprising the amino acid sequences of SEQ ID NO: 2, 5, 7, or 10. The nucleic acid can be, for example, a genomic DNA fragment or a cDNA molecule that includes the nucleic acid sequence of any of SEQ ID NO: 1, 3, 4, 6, 8, or 9.

Also included in the invention is an oligonucleotide, *e.g.*, an oligonucleotide which includes at least 6 contiguous nucleotides of a NOVX nucleic acid (*e.g.*, SEQ ID NO: 1, 3, 4, 6, 8, or 9) or a complement of said oligonucleotide. Also included in the invention are substantially purified NOVX polypeptides (SEQ ID NO: 2, 5, 7, or 10). In certain embodiments, the NOVX polypeptides include an amino acid sequence that is substantially identical to the amino acid sequence of a human NOVX polypeptide.

The invention also features antibodies that immunoselectively bind to NOVX polypeptides, or fragments, homologs, analogs or derivatives thereof.

In another aspect, the invention includes pharmaceutical compositions that include therapeutically- or prophylactically-effective amounts of a therapeutic and a pharmaceutically-acceptable carrier. The therapeutic can be, *e.g.*, a NOVX nucleic acid, a NOVX polypeptide, or an antibody specific for a NOVX polypeptide. In a further aspect, the invention includes, in one or more containers, a therapeutically- or prophylactically-effective amount of this pharmaceutical composition.

In a further aspect, the invention includes a method of producing a polypeptide by culturing a cell that includes a NOVX nucleic acid, under conditions allowing for expression of the NOVX polypeptide encoded by the DNA. If desired, the NOVX polypeptide can then be recovered.

In another aspect, the invention includes a method of detecting the presence of a NOVX polypeptide in a sample. In the method, a sample is contacted with a compound that selectively binds to the polypeptide under conditions allowing for formation of a complex between the polypeptide and the compound. The complex is detected, if present, thereby identifying the NOVX polypeptide within the sample.

The invention also includes methods to identify specific cell or tissue types based on their expression of a NOVX.

Also included in the invention is a method of detecting the presence of a NOVX nucleic acid molecule in a sample by contacting the sample with a NOVX nucleic acid probe or primer,

and detecting whether the nucleic acid probe or primer bound to a NOVX nucleic acid molecule in the sample.

In a further aspect, the invention provides a method for modulating the activity of a NOVX polypeptide by contacting a cell sample that includes the NOVX polypeptide with a compound that binds to the NOVX polypeptide in an amount sufficient to modulate the activity of said polypeptide. The compound can be, *e.g.*, a small molecule, such as a nucleic acid, peptide, polypeptide, peptidomimetic, carbohydrate, lipid or other organic (carbon containing) or inorganic molecule, as further described herein.

Also within the scope of the invention is the use of a therapeutic in the manufacture of a medicament for treating or preventing disorders or syndromes including, *e.g.*, Cancer, Leukodystrophies, Breast cancer, Ovarian cancer, Prostate cancer, Uterine cancer, Hodgkin disease, Adenocarcinoma, Adrenoleukodystrophy,Cystitis, incontinence, Von Hippel-Lindau (VHL) syndrome, hypercalceimia, Endometriosis, Hirschsprung's disease, Crohn's Disease, Appendicitis, Cirrhosis, Liver failure, Wolfram Syndrome, Smith-Lemli-Opitz syndrome, Retinitis pigmentosa, Leigh syndrome; Congenital Adrenal Hyperplasia, Xerostomia; tooth decay and other dental problems; Inflammatory bowel disease, Diverticular disease, fertility, Infertility, cardiomyopathy, atherosclerosis, hypertension, congenital heart defects, aortic stenosis , atrial septal defect (ASD), atrioventricular (A-V) canal defect, ductus arteriosus, pulmonary stenosis , subaortic stenosis, ventricular septal defect (VSD), valve diseases, tuberous sclerosis, scleroderma, Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, obesity, Diabetes Insipidus and Mellitus with Optic Atrophy and Deafness, Pancreatitis, Metabolic Dysregulation, transplantation recovery, Autoimmune disease, Systemic lupus erythematosus, asthma, arthritis, psoriasis, Emphysema, Scleroderma, allergy, ARDS, Immunodeficiencies, Graft vesus host, Alzheimer's disease, Stroke, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Multiple sclerosis,Ataxia-telangiectasia, Behavioral disorders, Addiction, Anxiety, Pain, Neurodegeneration, Muscular dystrophy,Lesch-Nyhan syndrome,Myasthenia gravis, schizophrenia, and other dopamine-dysfunctional states, levodopa-induced dyskinesias, alcoholism, pileptic seizures and other neurological disorders, mental depression, Cerebellar ataxia, pure; Episodic ataxia, type 2; Hemiplegic migraine, Spinocerebellar ataxia-6, Tuberous sclerosis, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Renal tubular acidosis, IgA nephropathy, and other pathologies and disorders of the like.

The therapeutic can be, *e.g.*, a NOVX nucleic acid, a NOVX polypeptide, or a NOVX-specific antibody, or biologically-active derivatives or fragments thereof.

For example, the compositions of the present invention will have efficacy for treatment of patients suffering from the diseases and disorders disclosed above and other pathologies and disorders of the like. The polypeptides can be used as immunogens to produce antibodies specific for the invention, and as vaccines. They can also be used to screen for potential agonist and antagonist compounds. For example, a cDNA encoding NOVX may be useful in gene therapy, and NOVX may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from the diseases and disorders disclosed above and other pathologies and disorders of the like.

The invention further includes a method for screening for a modulator of disorders or syndromes including, *e.g.*, the diseases and disorders disclosed above and other pathologies and disorders of the like. The method includes contacting a test compound with a NOVX polypeptide and determining if the test compound binds to said NOVX polypeptide. Binding of the test compound to the NOVX polypeptide indicates the test compound is a modulator of activity, or of latency or predisposition to the aforementioned disorders or syndromes.

Also within the scope of the invention is a method for screening for a modulator of activity, or of latency or predisposition to disorders or syndromes including, *e.g.*, the diseases and disorders disclosed above and other pathologies and disorders of the like by administering a test compound to a test animal at increased risk for the aforementioned disorders or syndromes. The test animal expresses a recombinant polypeptide encoded by a NOVX nucleic acid. Expression or activity of NOVX polypeptide is then measured in the test animal, as is expression or activity of the protein in a control animal which recombinantly-expresses NOVX polypeptide and is not at increased risk for the disorder or syndrome. Next, the expression of NOVX polypeptide in both the test animal and the control animal is compared. A change in the activity of NOVX polypeptide in the test animal relative to the control animal indicates the test compound is a modulator of latency of the disorder or syndrome.

In yet another aspect, the invention includes a method for determining the presence of or predisposition to a disease associated with altered levels of a NOVX polypeptide, a NOVX nucleic acid, or both, in a subject (*e.g.*, a human subject). The method includes measuring the amount of the NOVX polypeptide in a test sample from the subject and comparing the amount of the polypeptide in the test sample to the amount of the NOVX polypeptide present in a control sample. An alteration in the level of the NOVX polypeptide in the test sample as compared to the control sample indicates the presence of or predisposition to a disease in the subject. Preferably, the predisposition includes, *e.g.*, the diseases and disorders disclosed above

and other pathologies and disorders of the like. Also, the expression levels of the new polypeptides of the invention can be used in a method to screen for various cancers as well as to determine the stage of cancers.

In a further aspect, the invention includes a method of treating or preventing a pathological condition associated with a disorder in a mammal by administering to the subject a NOVX polypeptide, a NOVX nucleic acid, or a NOVX-specific antibody to a subject (*e.g.*, a human subject), in an amount sufficient to alleviate or prevent the pathological condition. In preferred embodiments, the disorder, includes, *e.g.*, the diseases and disorders disclosed above and other pathologies and disorders of the like.

In yet another aspect, the invention can be used in a method to identity the cellular receptors and downstream effectors of the invention by any one of a number of techniques commonly employed in the art. These include but are not limited to the two-hybrid system, affinity purification, co-precipitation with antibodies or other specific-interacting molecules.

NOVX nucleic acids and polypeptides are further useful in the generation of antibodies that bind immuno-specifically to the novel NOVX substances for use in therapeutic or diagnostic methods. These NOVX antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-NOVX Antibodies" section below. The disclosed NOVX proteins have multiple hydrophilic regions, each of which can be used as an immunogen. These NOVX proteins can be used in assay systems for functional analysis of various human disorders, which will help in understanding of pathology of the disease and development of new drug targets for various disorders.

The NOVX nucleic acids and proteins identified here may be useful in potential therapeutic applications implicated in (but not limited to) various pathologies and disorders as indicated below. The potential therapeutic applications for this invention include, but are not limited to: protein therapeutic, small molecule drug target, antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), diagnostic and prognostic marker, gene therapy (gene delivery/gene ablation), research tools, tissue regeneration in vivo and in vitro of all tissues and cell types composing (but not limited to) those defined here.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present

specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides novel nucleotides and polypeptides encoded thereby. Included in the invention are the novel nucleic acid sequences and their encoded polypeptides. The sequences are collectively referred to herein as "NOVX nucleic acids" or "NOVX polynucleotides" and the corresponding encoded polypeptides are referred to as "NOVX polypeptides" or "NOVX proteins." Unless indicated otherwise, "NOVX" is meant to refer to any of the novel sequences disclosed herein. Table A provides a summary of the NOVX nucleic acids and their encoded polypeptides.

TABLE A. Sequences and Corresponding SEQ ID Numbers

NOVX No.	Internal Acc. No.	Nucleic Acid SEQ ID NO.	Polypeptide SEQ ID NO.
1a	20936375_0_1_da2	1	2
1b	20936375_0_1_da2_13374173	3	2
1c	20936375_0_1_da2_13374284	4	5
1d	20936375_0_1_da2_13373889	6	7
1e	20936375_0_1_da2_13374169	8	2
2	CG51523-04	9	10

NOVX nucleic acids and their encoded polypeptides are useful in a variety of applications and contexts. The various NOVX nucleic acids and polypeptides according to the invention are useful as novel members of the protein families according to the presence of domains and sequence relatedness to previously described proteins. Additionally, NOVX nucleic acids and polypeptides can also be used to identify proteins that are members of the family to which the NOVX polypeptides belong.

The present invention is based in part on nucleic acids encoding proteins that are new members of the endozepine-related protein family. Thus, the NOVX nucleic acids, polypeptides, antibodies, and related compounds according to the invention will be useful in therapeutic and diagnostic applications implicated in, for example: Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, stroke, tuberous sclerosis, hypercalceimia, Parkinson's disease, Huntington's disease, cerebral palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis,

ataxia-telangiectasia, leukodystrophies, behavioral disorders, addiction, anxiety, pain, neuroprotection, anxiety, depression, stress, immune dysfunction, alcoholism, obesity and diabetes and other pathologies/disorders.

The NOVX nucleic acids and polypeptides can also be used to screen for molecules, which inhibit or enhance NOVX activity or function. Specifically, the nucleic acids and polypeptides according to the invention may be used as targets for the identification of small molecules that modulate or inhibit, *e.g.*, neurogenesis, cell differentiation, cell proliferation, hematopoiesis, wound healing and angiogenesis.

Additional utilities for the NOVX nucleic acids and polypeptides according to the invention are disclosed herein.

NOV1

NOV1 includes a novel endozepine-related precursor-like protein and 4 variants. The disclosed sequences have been named NOV1a-e. A multiple sequence alignment highlighting the differing amino acids is shown below in Table 6.

NOV1a

NOV1a includes a novel endozepine-related protein disclosed below. A disclosed NOV1 nucleic acid of 3887 nucleotides (also referred to as 20936375_0_1_da2) encoding a novel endozepine-related protein is shown in Table 1A. An open reading frame was identified beginning with an ATG initiation codon at nucleotides 67-69 and ending with a TGA codon at nucleotides 1636-1638. A putative untranslated region upstream from the initiation codon and downstream from the termination codon is underlined in Table 1A. The start and stop codons are in bold letters.

Table 1A. NOV1a nucleotide sequence (SEQ ID NO:1).

GGGAACTGACCTGCTTAGTTCCCGGGCCTCCTCCTTTTGGGGCATGTTGATCCGCGGCTGCGCTCC**ATG**TTC
CAGTTTCATGCAGGCTCTTGGGAAAGCTGGTGCTGCTGCTGCCTGATTCCCGCCGACAGACCTTGGGACCGG
GGCCAACACTGGCAGCTGGAGATGGCGGACACGAGATCCGTGCACGAGACTAGGTTTGAGGCGGCCGTGAAG
GTGATCCAGAGTTTGCCGAAGAATGGTTCATTCCAGCCAACAAATGAAATGATGCTTAAATTTTATAGCTTC
TATAAGCAGGCAACTGAAGGACCCTGTAAACTTTCAAGGCCTGGATTTTGGGATCCTATTGGAAGATATAAA
TGGGATGCTTGGAGTTCACTGGGTGATATGACCAAAGAGGAAGCCATGATTGCATATGTTGAAGAAATGAAA
AAGATTATTGAAACTATGCCAATGACTGAGAAAGTTGAAGAATTGCTGCGTGTCATAGGTCCATTTTATGAA
ATTGTCGAGGACAAAAAGAGTGGCAGGAGTTCTGATATAACCTCAGATCTTGGTAATGTTCTCACTTCTGCT
CCGAACGCCAAAACCGTTAATGGTAAAGCTGAAAGCAGTGACAGTGGAGCCGAGTCTGAGGAAGAAGAGGCC
CAAGAAGAAGTGAAAGGAGCAGAACAAAGTGATAATGATAAGAAAATGATGAAGAAGTCAGCAGACCATAAG
AATTTGGAAGTCATTGTCACTAATGGCTATGATAAAGATGGCTTTGTTCAGGATATACAGAATGACATTCAT
GCCAGTTCTTCCCTGAATGGCAGAAGCACTGAAGAAGTAAAGCCCATTGATGAAAACTTGGGGCAAACTGGA
AAATCTGCTGTTTGCATTCACCAAGATATAAATGATGATCATGTTGAAGATGTTACAGGAATTCAGCATTTG
ACAAGCGATTCAGACAGTGAAGTTTACTGTGATTCTATGGAACAATTTGGACAAGAAGAGTCTTTAGACAGC
TTTACGTCCAACAATGGACCATTTCAGTATTACTTGGGTGGTCATTCCAGTCAACCCATGGAAAATTCTGGA
TTTCGTGAAGATATTCAAGTACCTCCTGGAAATGGCAACATTGGGAATATGCAGGTGGTTGCAGTTGAAGGA
AAAGGTGAAGTCAAGCATGGAGGAGAAGATGGCAGGAATAACAGCGGAGCACCACACCGGGAGAAGCGAGGC

GGAGAAACTGACGAATTCTCTAATGTTAGAAGAGGAAGAGGACATAGGATACAACACTTGAGCGAAGGAACC
AAGGGCCGGCAGGTGGGAAGTGGAGGTGATGGGGAGCGCTGGGGCTCCGACAGAGGGTCCCGAGGCAGCCTC
AATGAGCAGATCGCCCTCGTGCTGATGAGACTGCAGGAGGACATGCAGAATGTCCTTCAGAGACTGCAGAAA
CTGGAAACGCTGACTGCTTTGCAGGCAAAATCATCAACATCAACATTGCAGACTGCTCCTCAGCCCACCTCA
CAGAGACCATCTTGGTGGCCCTTCGAGATGTCTCCTGGTGTGCTAACGTTTGCCATCATATGGCCTTTTATT
GCACAGTGGTTGGTGTATTTATACTATCAAAGAAGGAGAAGAAAACTGAACTGAGGAAAATGGTGTTTTCCT
CAAGAAGACTACTGGAACTGGATGACCTCAGAATGAACTGGATTGTGGTGTTCACAAGAAAATCTTAGTTTG
TGATGATTACATTGCTTTTTGTTGTCCAGTAGTTTAGTTTGTGTACATATATACACATATATATTTTGCACT
ACACAAACGATAACATTTTAAGGACTAATATTGCTGATACTTGAATAATCAATCTCTACTAGGTTATAAGTA
GTATACACAGATTTACCCTGCCCTTGAACTTGAAGGACATTAAATTATTAATGATCATTTGGTAACATGTTT
ACCTGATTATCTTCCATAGAGTAACATAAGCTGCTTTTCAAAGGTACCATTGTGATAATGAGATCAAATTTA
TAAGTTATTATTTTTAATTTTCTAAATTAAATAAAAGAAAGAATGCAAACCAGGAGTGAATTTCAAATGAGA
TGTAATCGACTTTATATCTTAGTCACGGAGTTGCCATGGCATGTAGTAGAAAACCACAGGAAGAATGGTCAT
ATTCACTTTGTGGGCTGCCCATAATCTTTCTTGGGCATTCACAACTCTTGAGTTTGGTGTTCAGGCATCATT
ATTAAAAAGTGGAGTCCTATGTACCAGACTGAGTTTTTACAAATGATTTGCAGGCTAGACATAACCCACTGA
TGGAAATGGTGAAGAATGAGCTTCATGTAGGTTTAAAAGTGTATTCTGAGCCTGTAGATGATTAATCAGGTT
TTTATTCAATCATATAAATGTATTCCTTTGTAATCATTTTTGTTTAACTGAGGATATCTAGTGTCTGCTTCA
TAGGGTGCTTTGAAATATAAAATGAAAACTTATTTATACTGTTTTTACAACAGTCAAAAAGGAAACACACGA
AAATCACTTTTCTGCAACTGATGAACTATATAGACTGGACTCTTAACCTCTTAGTGCCTCAGTTTTTCCTTC
GGGGTATACTATTTTAGATACACCTACTTCACAGGAGTACTGGAGGGATTTGCAAGCTAACGGGCCAAATGC
TTCCATAGGAATAAGGCATGCCCAGCTAACAGAAATTTAAGTCCCTCTTCCCCACCTCTCTCATCTAGACCA
AAAAGAAGACTACAATTTAGATCCTTGGAGACTTTTCCTGTATGCCTTCCAAAACTTCCTGCTCACTTAGAT
GTACCTTGTGCTTTGAAGATTTCTTTGATGCAGTCTTTGTAACAAATAATTTTTTTTTTTAAATGGAGTTTC
TCTCTTGTTGCCCAGGCTGGAGGGCAATGGCGTGATCTTGGCTCACTGCAACTTCCACCTCCTGGGTTCAAG
TGATTCTCCTGCCTCACTCCCAAGTAGCTGGGATTACAGGCATGCACCACCATGCCTGGCTAATTTTGTATT
TTTAGTAGAGACGGGGTTTCTCCATGTTGGTAAGGCTGATCTCAAACTCCCAACCTCAGGTGATCTGCCTGC
CTCGGCCTCCCAAAGTGCTGGGATTAGAGGCGTGAGCCACTGCACCTGGCCTGTAGCAAATAATTTTTAAGC
ATTCTCAAGATGTATGATGTTGGGTTTAACATCATATGTTCACAGTGTTTTAAATAAGAAATAATCTGTCTT
TAGTATACAGGATGGTTCTTTGTTTGCCTGAAAAGTATAAGAATACAATTTACTTTTCCCAAACTCTTTTTC
CTTATTTTTTTCTTTCAAATAACTTTAATTATTTAATCCCATACTGATTAAAATATGTCTGTCTAAAGGGAT
CTACTACTATTTGCTTTTAAAAAAGTGTTCCCTATTTATTTAGGAAAAAGTGAAGCAAGCAACTGAATTTAT
ATGTAAAAATAACATTTAGACCTGTGGATCAAAGATTATTTTCAAAAATGAGATATCAATTTGGTATCAGTA
AAAAGTGTCTTTTCTCTTAGTATATAAATAAATAAAATGCAAATTAAGAGTGTAAGCAGTTGTGAATCATTA
CAAGCAGTCTGGGATAACTAACTCCTAGGGATAACCATTTCTATCCTGTTCCAAAGTCTTATTTTATAGTTT
GAAAAGCACTTTGCACACAGTTCTTGTATATAAAAGTAAAGATGTAATTATAGGATATAGTGTTCCTGCTTT
GTTTAATAAGAACCTCATTTAAACTTGACAGCTATGGTATTTTTTTAATGATAACTTCTTTTCTGTTTATTG
TAAAACTAAGTTAAAAATAAAAGGTTAATGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO:1)

In all BLAST alignments herein, the "E-value" or "Expect" value is a numeric indication of the probability that the aligned sequences could have achieved their similarity to the BLAST query sequence by chance alone, within the database that was searched. For example, the probability that the subject ("Sbjct") retrieved from the NOV1 BLAST analysis, e.g., human CAC24877 mRNA, matched the Query NOV1 sequence purely by chance is 3.6e-284. The Expect value (E) is a parameter that describes the number of hits one can "expect" to see just by chance when searching a database of a particular size. It decreases exponentially with the Score (S) that is assigned to a match between two sequences. Essentially, the E value describes the random background noise that exists for matches between sequences.

The Expect value is used as a convenient way to create a significance threshold for reporting results. The default value used for blasting is typically set to 0.0001. In BLAST 2.0, the Expect value is also used instead of the P value (probability) to report the significance of

matches. For example, an E value of one assigned to a hit can be interpreted as meaning that in a database of the current size one might expect to see one match with a similar score simply by chance. An E value of zero means that one would not expect to see any matches with a similar score simply by chance. See, e.g., http://www.ncbi.nlm.nih.gov/Education/BLASTinfo/.

Occasionally, a string of X's or N's will result from a BLAST search. This is a result of automatic filtering of the query for low-complexity sequence that is performed to prevent artifactual hits. The filter substitutes any low-complexity sequence that it finds with the letter "N" in nucleotide sequence (e.g., "NNNNNNNNNNNN") or the letter "X" in protein sequences (e.g., "XXXXXXXXX"). Low-complexity regions can result in high scores that reflect compositional bias rather than significant position-by-position alignment. (Wootton and Federhen, Methods Enzymol 266:554-571, 1996).

The disclosed NOV1a polypeptide (SEQ ID NO:2) encoded by SEQ ID NO:1 has 523 amino acid residues and is presented in Table 2B using the one-letter amino acid code. Signal P, Psort and Hydropathy results predict that NOV1a has a signal peptide and is likely to be localized in the plasma membrane with a certainty of 0.7300.

Laboratory cloning of NOV1a through SeqCalling™ data is described in Example 1.

Table 1B. Encoded NOV1a protein sequence (SEQ ID NO:2).

```
MFQFHAGSWESWCCCCLIPADRPWDRGQHWQLEMADTRSVHETRFEAAVKVIQSLPKNGSFQPTNEMMLK
FYSFYKQATEGPCKLSRPGFWDPIGRYKWDAWSSLGDMTKEEAMIAYVEEMKKIIETMPMTEKVEELLRV
IGPFYEIVEDKKSGRSSDITSDLGNVLTSAPNAKTVNGKAESSDSGAESEEEEAQEEVKGAEQSDNDKKM
MKKSADHKNLEVIVTNGYDKDGFVQDIQNDIHASSSLNGRSTEEVKPIDENLGQTGKSAVCIHQDINDDH
VEDVTGIQHLTSDSDSEVYCDSMEQFGQEESLDSFTSNNGPFQYYLGGHSSQPMENSGFREDIQVPPGNG
NIGNMQVVAVEGKGEVKHGGEDGRNNSGAPHREKRGGETDEFSNVRRGRGHRIQHLSEGTKGRQVGSGGD
GERWGSDRGSRGSLNEQIALVLMRLQEDMQNVLQRLQKLETLTALQAKSSTSTLQTAPQPTSQRPSWWPF
EMSPGVLTFAIIWPFIAQWLVYLYYQRRRRKLN (SEQ ID NO:2)
```

The full amino acid sequence of the disclosed NOV1a protein was found to have 518 of 534 amino acid residues (97%) identical to, and 520 of 534 amino acid residues (97%) similar to, the 534 amino acid residue ptnr:REMTRMBL-ACC:CAC24877 protein from sequence 23 of patent WO0078802. Public amino acid databases include the GenBank databases, SwissProt, PDB and PIR.

NOV1a maps to chromosome 10 and is expressed in at least the following tissues: : Brain, Colon, Foreskin, Kidney, Larynx, Lung, Mammary gland/Breast, Ovary, Pancreas, Placenta, Retina, Small Intestine, Spleen, Testis, Thalamus, and Uterus.

The amino acid sequence of NOV1a had high homology to other proteins as shown in Table 1C.

Table 1C. BLASTX results for NOV1a		
Sequences producing High-scoring Segment Pairs:	High Score	Smallest Sum Prob P(N)
patp:AAB48379 Human SEC12 protein sequence (clone ID 2093...	2733	2.6e-284
patp:AAU00399 Human secreted protein, POLY11 - Homo sapie...	2733	2.6e-284
patp:AAB48375 Human SEC8 protein sequence (clone ID 20936...	2727	1.1e-283
patp:AAB81816 Human endozepine-like ENDO6 SEQ ID NO: 23 -...	2687	1.9e-279
patp:AAB81816 Human endozepine-like ENDO6 SEQ ID NO: 23 -...	2687	1.9e-279

The disclosed NOV1a polypeptide also has homology to the amino acid sequences shown in the BLASTP data listed in Table 1D.

Table 1D. BLAST results for NOV1a					
Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
CAC24877	Sequence 23 from Patent WO0078802/human	534	518/534 (97%)	520/534 (97%)	3.6e-284
CAC24873	Sequence 15 from Patent WO0078802/human	536	517/531 (97%)	518/531 (97%)	1.6e-283
P07106	Endozepine-related protein precursor/bovine	533	443/533 (83%)	473/533 (88%)	9.8e-243
Q9CW41	1300014E15RIK Protein	504	389/517 (75%)	433/517 (83%)	5.9e-197
Q9UFB5	Hypothetical 31.5kDa Protein/human	283	282/283 (99%)	283/283 (100%)	3.4e-153

The presence of identifiable domains in NOV1a was determined by searches using software algorithms such as PROSITE, DOMAIN, Blocks, Pfam, ProDomain, and Prints, and then determining the Interpro number by crossing the domain match (or numbers) using the Interpro website (http:www.ebi.ac.uk/ interpro). DOMAIN results for NOV1a as disclosed in Tables 1E, were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST analyses. This BLAST analysis software samples domains found in theSmart and Pfam collections. For Table 1E and all successive DOMAIN sequence alignments, fully conserved single residues are indicated by black shading or by the sign (|) and "strong" semi-conserved residues are indicated by grey shading or by the sign (+). The "strong" group of conserved amino acid residues may be any one of the following groups of amino acids: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW.

Table 1E lists the domain description from DOMAIN analysis results against NOV1a. This indicates that the NOV1a sequence has properties similar to those of other proteins known to contain this domain.

Table 1E. Domain Analysis of NOV1a

```
ACBP (InterPro)  Acyl CoA binding protein

ACBP: domain 1 of 1, from 41 to 129: score 199.7, E = 4.4e-56
                    *->lqedFeaAaekvkkLkknGpvkPSneekLkLYsLYKQATvGDvnter
                       ++++FeaA++++++L+knG+++P+ne++Lk+Ys+YKQAT+G+++++r
  NOV1a          41    HETRFEAAVKVIQSLPKNGSFQPTNEMMLKFYSFYKQATEGPCKLSR 87

                    PGmfDlkgrAKWDAWnelkGmSkeeAmkaYIakVeeLiakya<-* (SEQ ID NO: 11)
                    PG++D++gr+KWDAW++l++m+keeAm+aY+++++++i++++
  NOV1a          88 PGFWDPIGRYKWDAWSSLGDMTKEEAMIAYVEEMKKIIETMP    129
```

NOV1b

In an alternative embodiment, a NOV1 variant is NOV1b of 3920 nucleotides (also referred to as 20936375_0_1_da1_13374173), shown in Table 2A. A NOV1b variant differs from NOVa by a single nucleotide polymorphism at position 612 of the nucleotide sequence. This change results in no change in the amino acid sequence.

Table 2A. NOV1b nucleotide sequence (SEQ ID NO: 3).

```
GGGAACTGACCTGCTTAGTTCCCGGGCCTCCTCCTTTTGGGGCATGTTGATCCGCGGCTGCGCTCCATGTTCCAGTTTCA
TGCAGGCTCTTGGGAAAGCTGGTGCTGCTGCTGCCTGATTCCCGCCGACAGACCTTGGGACCGGGGCCAACACTGGCAGC
TGGAGATGGCGGACACGAGATCCGTGCACGAGACTAGGTTTGAGGCGGCCGTGAAGGTGATCCAGAGTTTGCCGAAGAAT
GGTTCATTCCAGCCAACAAATGAAATGATGCTTAAATTTTATAGCTTCTATAAGCAGGCAACTGAAGGACCCTGTAAACT
TTCAAGGCCTGGATTTTGGGATCCTATTGGAAGATATAAATGGGATGCTTGGAGTTCACTGGGTGATATGACCAAAGAGG
AAGCCATGATTGCATATGTTGAAGAAATGAAAAAGATTATTGAAACTATGCCAATGACTGAGAAAGTTGAAGAATTGCTG
CGTGTCATAGGTCCATTTTATGAAATTGTCGAGGACAAAAAGAGTGGCAGGAGTTCTGATATAACCTCAGTCCGACTGGA
GAAAATCTCTAAATGTTTAGAAGATCTTGGTAATGTTCTCACTTCTGCTCCAAACGCCAAAACCGTTAATGGTAAAGCTG
AAAGCAGTGACAGTGGAGCCGAGTCTGAGGAAGAAGAGGCCCAAGAAGAAGTGAAAGGAGCAGAACAAAGTGATAATGAT
AAGAAAATGATGAAGAAGTCAGCAGACCATAAGAATTTGGAAGTCATTGTCACTAATGGCTATGATAAAGATGGCTTTGT
TCAGGATATACAGAATGACATTCATGCCAGTTCTTCCCTGAATGGCAGAAGCACTGAAGAAGTAAAGCCCATTGATGAAA
ACTTGGGGCAAACTGGAAAATCTGCTGTTTGCATTCACCAAGATATAAATGATGATCATGTTGAAGATGTTACAGGAATT
CAGCATTTGACAAGCGATTCAGACAGTGAAGTTTACTGTGATTCTATGGAACAATTTGGACAAGAAGAGTCTTTAGACAG
CTTTACGTCCAACAATGGACCATTTCAGTATTACTTGGGTGGTCATTCCAGTCAACCCATGGAAAATTCTGGATTTCGTG
AAGATATTCAAGTACCTCCTGGAAATGGCAACATTGGGAATATGCAGGTGGTTGCAGTTGAAGGAAAAGGTGAAGTCAAG
CATGGAGGAGAAGATGGCAGGAATAACAGCGGAGCACCACACCGGGAGAAGCGAGGCGGAGAAACTGACGAATTCTCTAA
TGTTAGAAGAGGAAGAGGACATAGGATACAACACTTGAGCGAAGGAACCAAGGGCCGGCAGGTGGGAAGTGGAGGTGATG
GGGAGCGCTGGGGCTCCGACAGAGGGTCCCGAGGCAGCCTCAATGAGCAGATCGCCCTCGTGCTGATGAGACTGCAGGAG
GACATGCAGAATGTCCTTCAGAGACTGCAGAAACTGGAAACGCTGACTGCTTTGCAGGCAAAATCATCAACATCAACATT
GCAGACTGCTCCTCAGCCCACCTCACAGAGACCATCTTGGTGGCCCTTCGAGATGTCTCCTGGTGTGCTAACGTTTGCCA
TCATATGGCCTTTTATTGCACAGTGGTTGGTGTATTTATACTATCAAAGAAGGAGAAGAAAACTGAACTGAGGAAAATGG
TGTTTTCCTCAAGAAGACTACTGGAACTGGATGACCTCAGAATGAACTGGATTGTGGTGTTCACAAGAAAATCTTAGTTT
GTGATGATTACATTGCTTTTTGTTGTCCAGTAGTTTAGTTTGTGTACATATATACACATATATATTTTGCACTACACAAA
CGATAACATTTTAAGGACTAATATTGCTGATACTTGAATAATCAATCTCTACTAGGTTATAAGTAGTATACACAGATTTA
CCCTGCCCTTGAACTTGAAGGACATTAAATTATTAATGATCATTTGGTAACATGTTTACCTGATTATCTTCCATAGAGTA
ACATAAGCTGCTTTTCAAAGGTACCATTGTGATAATGAGATCAAATTTATAAGTTATTATTTTTAATTTTCTAAATTAAA
TAAAAGAAAGAATGCAAACCAGGAGTGAATTTCAAATGAGATGTAATCGACTTTATATCTTAGTCACGGAGTTGCCATGG
CATGTAGTAGAAAACCACAGGAAGAATGGTCATATTCACTTTGTGGGCTGCCCATAATCTTTCTTGGGCATTCACAACTC
TTGAGTTTGGTGTTCAGGCATCATTATTAAAAAGTGGAGTCCTATGTACCAGACTGAGTTTTTACAAATGATTTGCAGGC
TAGACATAACCCACTGATGGAAATGGTGAAGAATGAGCTTCATGTAGGTTTAAAAGTGTATTCTGAGCCTGTAGATGATT
AATCAGGTTTTTATTCAATCATATAAATGTATTCCTTTGTAATCATTTTTGTTTAACTGAGGATATCTAGTGTCTGCTTC
ATAGGGTGCTTTGAAATATAAAATGAAAACTTATTTATACTGTTTTTACAACAGTCAAAAAGGAAACACACGAAAATCAC
TTTTCTGCAACTGATGAACTATATAGACTGGACTCTTAACCTCTTAGTGCCTCAGTTTTTCCTTCGGGGTATACTATTTT
AGATACACCTACTTCACAGGAGTACTGGAGGGATTTGCAAGCTAACGGGCCAAATGCTTCCATAGGAATAAGGCATGCCC
AGCTAACAGAAATTTAAGTCCCTCTTCCCCACCTCTCTCATCTAGACCAAAAAGAAGACTACAATTTAGATCCTTGGAGA
```

CTTTTCCTGTATGCCTTCCAAAACTTCCTGCTCACTTAGATGTACCTTGTGCTTTGAAGATTTCTTTGATGCAGTCTTTG
TAACAAATAATTTTTTTTTTTAAATGGAGTTTCTCTCTTGTTGCCCAGGCTGGAGGGCAATGGCGTGATCTTGGCTCACT
GCAACTTCCACCTCCTGGGTTCAAGTGATTCTCCTGCCTCACTCCCAAGTAGCTGGGATTACAGGCATGCACCACCACTGC
CTGGCTAATTTTGTATTTTTAGTAGAGACGGGGTTTCTCCATGTTGGTAAGGCTGATCTCAAACTCCCAACCTCAGGTGA
TCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTAGAGGCGTGAGCCACTGCACCTGGCCTGTAGCAAATAATTTTTAAG
CATTCTCAAGATGTATGATGTTGGGTTTAACATCATATGTTCACAGTGTTTTAAATAAGAAATAATCTGTCTTTAGTATA
CAGGATGGTTCTTTGTTTGCCTGAAAAGTATAAGAATACAATTTACTTTTCCCAAACTCTTTTTCCTTATTTTTTTCTTT
CAAATAACTTTAATTATTTAATCCCATACTGATTAAAATATGTCTGTCTAAAGGGATCTACTACTATTTGCTTTTAAAAA
AGTGTTCCCTATTTATTTAGGAAAAAGTGAAGCAAGCAACTGAATTTATATGTAAAAATAACATTTAGACCTGTGGATCA
AAGATTATTTTCAAAAATGAGATATCAATTTGGTATCAGTAAAAAGTGTCTTTTCTCTTAGTATATAAATAAATAAAATG
CAAATTAAGAGTGTAAGCAGTTGTGAATCATTACAAGCAGTCTGGGATAACTAACTCCTAGGGATAACCATTTCTATCCT
GTTCCAAAGTCTTATTTTATAGTTTGAAAAGCACTTTGCACACAGTTCTTGTATATAAAAGTAAAGATGTAATTATAGGA
TATAGTGTTCCTGCTTTGTTTAATAAGAACCTCATTTAAACTTGACAGCTATGGTATTTTTTTAATGATAACTTCTTTTC
TGTTTATTGTAAAACTAAGTTAAAAATAAAAGGTTAATGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 3)

NOV1c

In an alternative embodiment, a NOV1 variant is NOV1c of 3920 nucleotides (also referred to as 20936375_0_1_da1_13374284), shown in Table 3A. A NOV1c variant differs from NOVa by a single nucleotide polymorphism at position 1252 of the nucleotide sequence that results in a change at amino acid 396 of the protein sequence.

Table 3A. NOV1c nucleotide sequence (SEQ ID NO: 4).

GGGAACTGACCTGCTTAGTTCCCGGGCCTCCTCCTTTTGGGGCATGTTGATCCGCGGCTGCGCTCCATGTTCCAGTTTCA
TGCAGGCTCTTGGGAAAGCTGGTGCTGCTGCTGCCTGATTCCCGCCGACAGACCTTGGGACCGGGGCCAACACTGGCAGC
TGGAGATGGCGGACACGAGATCCGTGCACGAGACTAGGTTTGAGGCGGCCGTGAAGGTGATCCAGAGTTTGCCGAAGAAT
GGTTCATTCCAGCCAACAAATGAAATGATGCTTAAATTTTATAGCTTCTATAAGCAGGCAACTGAAGGACCCTGTAAACT
TTCAAGGCCTGGATTTGGGATCCTATTGGAAGATATAAATGGGATGCTTGGAGTTCACTGGGTGATATGACCAAAGAGG
AAGCCATGATTGCATATGTTGAAGAAATGAAAAAGATTATTGAAACTATGCCAATGACTGAGAAAGTTGAAGAATTGCTG
CGTGTCATAGGTCCATTTTATGAAATTGTCGAGGACAAAAAGAGTGGCAGGAGTTCTGATATAACCTCAGTCCGACTGGA
GAAAATCTCTAAATGTTTAGAAGATCTTGGTAATGTTCTCACTTCTGCTCCGAACGCCAAAACCGTTAATGGTAAAGCTG
AAAGCAGTGACAGTGGAGCCGAGTCTGAGGAAGAAGAGGCCCAAGAAGAAGTGAAAGGAGCAGAACAAAGTGATAATGAT
AAGAAAATGATGAAGAAGTCAGCAGACCATAAGAATTTGGAAGTCATTGTCACTAATGGCTATGATAAAGATGGCTTTGT
TCAGGATATACAGAATGACATTCATGCCAGTTCTTCCCTGAATGGCAGAAGCACTGAAGAAGTAAAGCCCATTGATGAAA
ACTTGGGGCAAACTGGAAAATCTGCTGTTTGCATTCACCAAGATATAAATGATGATCATGTTGAAGATGTTACAGGAATT
CAGCATTTGACAAGCGATTCAGACAGTGAAGTTTACTGTGATTCTATGGAACAATTTGGACAAGAAGAGTCTTTAGACAG
CTTTACGTCCAACAATGGACCATTTCAGTATTACTTGGGTGGTCATTCCAGTCAACCCATGGAAAATTCTGGATTTCGTG
AAGATATTCAAGTACCTCCTGGAAATGGCAACATTGGGAATATGCAGGTGGTTGCAGTTGAAGGAAAAGGTGAAGTCAAG
CATGGAGGAGAAGATGGCAGGAATAACAGCGGAGCACCACACCGGGAGAAGTGAGGCGGAGAAACTGACGAATTCTCTAA
TGTTAGAAGAGGAAGAGGACATAGGATACAACACTTGAGCGAAGGAACCAAGGGCCGGCAGGTGGGAAGTGGAGGTGATG
GGGAGCGCTGGGGCTCCGACAGAGGGTCCCGAGGCAGCCTCAATGAGCAGATCGCCCTCGTGCTGATGAGACTGCAGGAG
GACATGCAGAATGTCCTTCAGAGACTGCAGAAACTGGAAACGCTGACTGCTTTGCAGGCAAAATCATCAACATCAACATT
GCAGACTGCTCCTCAGCCCACCTCACAGAGACCATCTTGGTGGCCCTTCGAGATGTCTCCTGGTGTGCTAACGTTTGCCA
TCATATGGCCTTTTATTGCACAGTGGTTGGTGTATTTATACTATCAAAGAAGGAGAAGAAAACTGAACTGAGGAAAATGG
TGTTTTCCTCAAGAAGACTACTGGAACTGGATGACCTCAGAATGAACTGGATTGTGGTGTTCACAAGAAAATCTTAGTTT
GTGATGATTACATTGCTTTTTGTTGTCCAGTAGTTTAGTTTGTGTACATATATACACATATATATTTTGCACTACACAAA
CGATAACATTTTAAGGACTAATATTGCTGATACTTGAATAATCAATCTCTACTAGGTTATAAGTAGTATACACAGATTTA
CCCTGCCCTTGAACTTGAAGGACATTAAATTATTAATGATCATTTGGTAACATGTTTACCTGATTATCTTCCATAGAGTA
ACATAAGCTGCTTTTCAAAGGTACCATTGTGATAATGAGATCAAATTTATAAGTTATTATTTTTAATTTTCTAAATTAAA
TAAAAGAAAGAATGCAAACCAGGAGTGAATTTCAAATGAGATGTAATCGACTTTATATCTTAGTCACGGAGTTGCCATGG
CATGTAGTAGAAAACCACAGGAAGAATGGTCATATTCACTTTGTGGGCTGCCCATAATCTTTCTTGGGCATTCACAACTC
TTGAGTTTGGTGTTCAGGCATCATTATTAAAAAGTGGAGTCCTATGTACCAGACTGAGTTTTTACAAATGATTTGCAGGC
TAGACATAACCCACTGATGGAAATGGTGAAGAATGAGCTTCATGTAGGTTTAAAAGTGTATTCTGAGCCTGTAGATGATT
AATCAGGTTTTTATTCAATCATATAAATGTATTCCTTTGTAATCATTTTTGTTTAACTGAGGATATCTAGTGTCTGCTTC
ATAGGGTGCTTTGAAATATAAAATGAAAACTTATTTATACTGTTTTTACAACAGTCAAAAAGGAAACACACGAAAATCAC
TTTTCTGCAACTGATGAACTATATAGACTGGACTCTTAACCTCTTAGTGCCTCAGTTTTTCCTTCGGGGTATACTATTTT
AGATACACCTACTTCACAGGAGTACTGGAGGGATTTGCAAGCTAACGGGCCAAATGCTTCCATAGGAATAAGGCATGCCC
AGCTAACAGAAATTTAAGTCCCTCTTCCCCACCTCTCTCATCTAGACCAAAAAGAAGACTACAATTTAGATCCTTGGAGA
CTTTTCCTGTATGCCTTCCAAAACTTCCTGCTCACTTAGATGTACCTTGTGCTTTGAAGATTTCTTTGATGCAGTCTTTG
TAACAAATAATTTTTTTTTTTAAATGGAGTTTCTCTCTTGTTGCCCAGGCTGGAGGGCAATGGCGTGATCTTGGCTCACT

GCAACTTCCACCTCCTGGGTTCAAGTGATTCTCCTGCCTCACTCCCAAGTAGCTGGGATTACAGGCATGCACCACCATGC
CTGGCTAATTTTGTATTTTTAGTAGAGACGGGGTTTCTCCATGTTGGTAAGGCTGATCTCAAACTCCCAACCTCAGGTGA
TCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTAGAGGCGTGAGCCACTGCACCTGGCCTGTAGCAAATAATTTTTAAG
CATTCTCAAGATGTATGATGTTGGGTTTAACATCATATGTTCACAGTGTTTTAAATAAGAAATAATCTGTCTTTAGTATA
CAGGATGGTTCTTTGTTTGCCTGAAAAGTATAAGAATACAATTTACTTTTCCCAAACTCTTTTTCCTTATTTTTTTCTTT
CAAATAACTTTAATTATTTAATCCCATACTGATTAAAATATGTCTGTCTAAAGGGATCTACTACTATTTGCTTTTAAAAA
AGTGTTCCCTATTTATTTAGGAAAAAAGTGAAGCAAGCAACTGAATTTATATGTAAAAATAACATTTAGACCTGTGGATCA
AAGATTATTTTCAAAAATGAGATATCAATTTGGTATCAGTAAAAAGTGTCTTTTCTCTTAGTATATAAATAAATAAAATG
CAAATTAAGAGTGTAAGCAGTTGTGAATCATTACAAGCAGTCTGGGATAACTAACTCCTAGGGATAACCATTTCTATCCT
GTTCCAAAGTCTTATTTTATAGTTTGAAAAGCACTTTGCACACAGTTCTTGTATATAAAAGTAAAGATGTAATTATAGGA
TATAGTGTTCCTGCTTTGTTTAATAAGAACCTCATTTAAACTTGACAGCTATGGTATTTTTTTAATGATAACTTCTTTTC
TGTTTATTGTAAAACTAAGTTAAAAATAAAAGGTTAATGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 4)

Table 3A. Encoded NOV1c protein sequence (SEQ ID NO: 5).

MFQFHAGSWESWCCCCLIPADRPWDRGQHWQLEMADTRSVHETRFEAAVKVIQSLPKNGSFQPTNEMMLKFYSFYKQATE
GPCKLSRPGFWDPIGRYKWDAWSSLGDMTKEEAMIAYVEEMKKIIETMPMTEKVEELLRVIGPFYEIVEDKKSGRSSDIT
SVRLEKISKCLEDLGNVLTSAPNAKTVNGKAESSDSGAESEEEEAQEEVKGAEQSDNDKKMMKKSADHKNLEVIVTNGYD
KDGFVQDIQNDIHASSSLNGRSTEEVKPIDENLGQTGKSAVCIHQDINDDHVEDVTGIQHLTSDSDSEVYCDSMEQFGQE
ESLDSFTSNNGPFQYYLGGHSSQPMENSGFREDIQVPPGNGNIGNMQVVAVEGKGEVKHGGEDGRNNSGAPHREK*GGET
DEFSNVRRGRGHRIQHLSEGTKGRQVGSGGDGERWGSDRGSRGSLNEQIALVLMRLQEDMQNVLQRLQKLETLTALQAKS
STSTLQTAPQPTSQRPSWWPFEMSPGVLTFAIIWPFIAQWLVYLYYQRRRRKLN (SEQ ID NO: 5)

NOV1d

In an alternative embodiment, a NOV1 variant is NOV1d of 3920 nucleotides (also referred to as 20936375_0_1_da1_13373889), shown in Table 4A. A NOV1d variant differs from NOVa by a single nucleotide polymorphism at position 1308 of the nucleotide sequence that results in an Ile to Met change at amino acid 414 of the protein sequence.

Table 4A. NOV1d nucleotide sequence (SEQ ID NO: 6).

GGGAACTGACCTGCTTAGTTCCCGGGCCTCCTCCTTTTGGGGCATGTTGATCCGCGGCTGCGCTCCATGTTCCAGTTTCA
TGCAGGCTCTTGGGAAAGCTGGTGCTGCTGCTGCCTGATTCCCGCCGACAGACCTTGGGACCGGGGCCAACACTGGCAGC
TGGAGATGGCGGACACGAGATCCGTGCACGAGACTAGGTTTGAGGCGGCCGTGAAGGTGATCCAGAGTTTGCCGAAGAAT
GGTTCATTCCAGCCAACAAATGAAATGATGCTTAAATTTTATAGCTTCTATAAGCAGGCAACTGAAGGACCCTGTAAACT
TTCAAGGCCTGGATTTTGGGATCCTATTGGAAGATATAAATGGGATGCTTGGAGTTCACTGGGTGATATGACCAAAGAGG
AAGCCATGATTGCATATGTTGAAGAAATGAAAAAGATTATTGAAACTATGCCAATGACTGAGAAAGTTGAAGAATTGCTG
CGTGTCATAGGTCCATTTTATGAAATTGTCGAGGACAAAAAGAGTGGCAGGAGTTCTGATATAACCTCAGTCCGACTGGA
GAAAATCTCTAAATGTTTAGAAGATCTTGGTAATGTTCTCACTTCTGCTCCGAACGCCAAAACCGTTAATGGTAAAGCTG
AAAGCAGTGACAGTGGAGCCGAGTCTGAGGAAGAAGAGGCCCAAGAAGAAGTGAAAGGAGCAGAACAAAGTGATAATGAT
AAGAAAATGATGAAGAAGTCAGCAGACCATAAGAATTTGGAAGTCATTGTCACTAATGGCTATGATAAAGATGGCTTTGT
TCAGGATATACAGAATGACATTCATGCCAGTTCTTCCCTGAATGGCAGAAGCACTGAAGAAGTAAAGCCCATTGATGAAA
ACTTGGGGCAAACTGGAAAATCTGCTGTTTGCATTCACCAAGATATAAATGATGATCATGTTGAAGATGTTACAGGAATT
CAGCATTTGACAAGCGATTCAGACAGTGAAGTTTACTGTGATTCTATGGAACAATTTGGACAAGAAGAGTCTTTAGACAG
CTTTACGTCCAACAATGGACCATTTCAGTATTACTTGGGTGGTCATTCCAGTCAACCCATGGAAAATTCTGGATTTCGTG
AAGATATTCAAGTACCTCCTGGAAATGGCAACATTGGGAATATGCAGGTGGTTGCAGTTGAAGGAAAAGGTGAAGTCAAG
CATGGAGGAGAAGATGGCAGGAATAACAGCGGAGCACCACACCGGGAGAAGCGAGGCGGAGAAACTGACGAATTCTCTAA
TGTTAGAAGAGGAAGAGGACATAGGATGCAACACTTGAGCGAAGGAACCAAGGGCCGGCAGGTGGGAAGTGGAGGTGATG
GGGAGCGCTGGGGCTCCGACAGAGGGTCCCGAGGCAGCCTCAATGAGCAGATCGCCCTCGTGCTGATGAGACTGCAGGAG
GACATGCAGAATGTCCTTCAGAGACTGCAGAAACTGGAAACGCTGACTGCTTTGCAGGCAAAATCATCAACATCAACATT
GCAGACTGCTCCTCAGCCCACCTCACAGAGACCATCTTGGTGGCCCTTCGAGATGTCTCCTGGTGTGCTAACGTTTGCCA
TCATATGGCCTTTTATTGCACAGTGGTTGGTGTATTTATACTATCAAAGAAGGAGAAGAAAACTGAACTGAGGAAAATGG
TGTTTTCCTCAAGAAGACTACTGGAACTGGATGACCTCAGAATGAACTGGATTGTGGTGTTCACAAGAAAATCTTAGTTT
GTGATGATTACATTGCTTTTTGTTGTCCAGTAGTTTAGTTTGTGTACATATATACACATATATATTTTGCACTACACAAA
CGATAACATTTTAAGGACTAATATTGCTGATACTTGAATAATCAATCTCTACTAGGTTATAAGTAGTATACACAGATTTA
CCCTGCCCTTGAACTTGAAGGACATTAAATTATTAATGATCATTTGGTAACATGTTTACCTGATTATCTTCCATAGAGTA
ACATAAGCTGCTTTTCAAAGGTACCATTGTGATAATGAGATCAAATTTATAAGTTATTATTTTTAATTTTCTAAATTAAA
TAAAAGAAAGAATGCAAACCAGGAGTGAATTTCAAATGAGATGTAATCGACTTTATATCTTAGTCACGGAGTTGCCATGG
CATGTAGTAGAAAACCACAGGAAGAATGGTCATATTCACTTTGTGGGCTGCCCATAATCTTTCTTGGGCATTCACAACTC

TTGAGTTTGGTGTTCAGGCATCATTATTAAAAAGTGGAGTCCTATGTACCAGACTGAGTTTTTACAAATGATTTGCAGGC
TAGACATAACCCACTGATGGAAATGGTGAAGAATGAGCTTCATGTAGGTTTAAAAGTGTATTCTGAGCCTGTAGATGATT
AATCAGGTTTTTATTCAATCATATAAATGTATTCCTTTGTAATCATTTTTGTTTAACTGAGGATATCTAGTGTCTGCTTC
ATAGGGTGCTTTGAAATATAAAATGAAAACTTATTTATACTGTTTTTACAACAGTCAAAAAGGAAACACACGAAAATCAC
TTTTCTGCAACTGATGAACTATATAGACTGGACTCTTAACCTCTTAGTGCCTCAGTTTTTCCTTCGGGGTATACTATTTT
AGATACACCTACTTCACAGGAGTACTGGAGGGATTTGCAAGCTAACGGGCCAAATGCTTCCATAGGAATAAGGCATGCCC
AGCTAACAGAAATTTAAGTCCCTCTTCCCCACCTCTCTCATCTAGACCAAAAAGAAGACTACAATTTAGATCCTTGGAGA
CTTTTCCTGTATGCCTTCCAAAACTTCCTGCTCACTTAGATGTACCTTGTGCTTTGAAGATTTCTTTGATGCAGTCTTTG
TAACAAATAATTTTTTTTTTTAAATGGAGTTTCTCTCTTGTTGCCCAGGCTGGAGGGCAATGGCGTGATCTTGGCTCACT
GCAACTTCCACCTCCTGGGTTCAAGTGATTCTCCTGCCTCACTCCCAAGTAGCTGGGATTACAGGCATGCACCACCATGC
CTGGCTAATTTTGTATTTTTAGTAGAGACGGGGTTTCTCCATGTTGGTAAGGCTGATCTCAAACTCCCAACCTCAGGTGA
TCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCTGGCCTGTAGCAAATAATTTTTAAG
CATTCTCAAGATGTATGATGTTGGGTTTAACATCATATGTTCACAGTGTTTTAAATAAGAAATAATCTGTCTTTAGTATA
CAGGATGGTTCTTTGTTTGCCTGAAAAGTATAAGAATACAATTTACTTTTCCCAAACTCTTTTTCCTTATTTTTTTCTTT
CAAATAACTTTAATTATTTAATCCCATACTGATTAAAATATGTCTGTCTAAAGGGATCTACTACTATTTGCTTTTAAAAA
AGTGTTCCCTATTTATTTAGGAAAAAGTGAAGCAAGCAACTGAATTTATATGTAAAAATAACATTTAGACCTGTGGATCA
AAGATTATTTTCAAAAATGAGATATCAATTTGGTATCAGTAAAAAGTGTCTTTTCTCTTAGTATATAAATAAATAAAATG
CAAATTAAGAGTGTAAGCAGTTGTGAATCATTACAAGCAGTCTGGGATAACTAACTCCTAGGGATAACCATTTCTATCCT
GTTCCAAAGTCTTATTTTATAGTTTGAAAAGCACTTTGCACACAGTTCTTGTATATAAAAGTAAAGATGTAATTATAGGA
TATAGTGTTCCTGCTTTGTTTAATAAGAACCTCATTTAAACTTGACAGCTATGGTATTTTTTTAATGATAACTTCTTTTC
TGTTTATTGTAAAACTAAGTTAAAAATAAAAGGTTAATGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 6)

Table 4B. Encoded NOV1d protein sequence (SEQ ID NO: 7).

MFQFHAGSWESWCCCCLIPADRPWDRGQHWQLEMADTRSVHETRFEAAVKVIQSLPKNGSFQPTNEMMLKFYSFYKQATE
GPCKLSRPGFWDPIGRYKWDAWSSLGDMTKEEAMIAYVEEMKKIIETMPMTEKVEELLRVIGPFYEIVEDKKSGRSSDIT
SVRLEKISKCLEDLGNVLTSAPNAKTVNGKAESSDSGAESEEEEAQEEVKGAEQSDNDKKMMKKSADHKNLEVIVTNGYD
KDGFVQDIQNDIHASSSLNGRSTEEVKPIDENLGQTGKSAVCIHQDINDDHVEDVTGIQHLTSDSDSEVYCDSMEQFGQE
ESLDSFTSNNGPFQYYLGGHSSQPMENSGFREDIQVPPGNGNIGNMQVVAVEGKGEVKHGGEDGRNNSGAPHREKRGGET
DEFSNVRRGRGHRMQHLSEGTKGRQVGSGGDGERWGSDRGSRGSLNEQIALVLMRLQEDMQNVLQRLQKLETLTALQAKS
STSTLQTAPQPTSQRPSWWPFEMSPGVLTFAIIWPFIAQWLVYLYYQRRRRKLN (SEQ ID NO: 7)

NOV1e

In an alternative embodiment, a NOV1 variant is NOV1e of 3920 nucleotides (also referred to as 20936375_0_1_da1_13374169), shown in Table 5A. A NOV1e variant differs from NOVa by a single nucleotide polymorphism at position 1698 of the nucleotide sequence. This results in no change in the protein sequence.

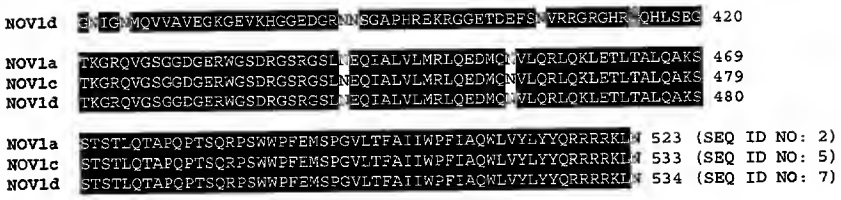
Table 5A. NOV1e nucleotide sequence (SEQ ID NO: 8).

GGGAACTGACCTGCTTAGTTCCCGGGCCTCCTCCTTTTGGGGCATGTTGATCCGCGGCTGCGCTCCATGTTCCAGTTTCA
TGCAGGCTCTTGGGAAAGCTGGTGCTGCTGCTGCCTGATTCCCGCCGACAGACCTTGGGACCGGGGCCAACACTGGCAGC
TGGAGATGGCGGACACGAGATCCGTGCACGAGACTAGGTTTGAGGCGGCCGTGAAGGTGATCCAGAGTTTGCCGAAGAAT
GGTTCATTCCAGCCAACAAATGAAATGATGCTTAAATTTTATAGCTTCTATAAGCAGGCAACTGAAGGACCCTGTAAACT
TTCAAGGCCTGGATTTTGGGATCCTATTGGAAGATATAAATGGGATGCTTGGAGTTCACTGGGTGATATGACCAAAGAGG
AAGCCATGATTGCATATGTTGAAGAAATGAAAAAGATTATTGAAACTATGCCAATGACTGAGAAAGTTGAAGAATTGCTG
CGTGTCATAGGTCCATTTTATGAAATTGTCGAGGACAAAAAGAGTGGCAGGAGTTCTGATATAACCTCAGTCCGACTGGA
GAAAATCTCTAAATGTTTAGAAGATCTTGGTAATGTTCTCACTTCTGCTCCGAACGCCAAAACCGTTAATGGTAAAGCTG
AAAGCAGTGACAGTGGAGCCGAGTCTGAGGAAGAAGAGGCCCAAGAAGAAGTGAAAGGAGCAGAACAAAGTGATAATGAT
AAGAAAATGATGAAGAAGTCAGCAGACCATAAGAATTTGGAAGTCATTGTCACTAATGGCTATGATAAAGATGGCTTTGT
TCAGGATATACAGAATGACATTCATGCCAGTTCTTCCCTGAATGGCAGAAGCACTGAAGAAGTAAAGCCCATTGATGAAA
ACTTGGGGCAAACTGGAAAATCTGCTGTTTGCATTCACCAAGATATAAATGATGATCATGTTGAAGATGTTACAGGAATT
CAGCATTTGACAAGCGATTCAGACAGTGAAGTTTACTGTGATTCTATGGAACAATTTGGACAAGAAGAGTCTTTAGACAG
CTTTACGTCCAACAATGGACCATTTCAGTATTACTTGGGTGGTCATTCCAGTCAACCCATGGAAAATTCTGGATTTCGTG
AAGATATTCAAGTACCTCCTGGAAATGGCAACATTGGGAATATGCAGGTGGTTGCAGTTGAAGGAAAAGGTGAAGTCAAG
CATGGAGGAGAAGATGGCAGGAATAACAGCGGAGCACCACACCGGGAGAAGCGAGGCGGAGAAACTGACGAATTCTCTAA
TGTTAGAAGAGGAAGAGGACATAGGATACAACACTTGAGCGAAGGAACCAAGGGCCGGCAGGTGGGAAGTGGAGGTGATG
GGGAGCGCTGGGGCTCCGACAGAGGGTCCCGAGGCAGCCTCAATGAGCAGATCGCCCTCGTGCTGATGAGACTGCAGGAG
GACATGCAGAATGTCCTTCAGAGACTGCAGAAACTGGAAACGCTGACTGCTTTGCAGGCAAAATCATCAACATCAACATT

GCAGACTGCTCCTCAGCCCACCTCACAGAGACCATCTTGGTGGCCCTTCGAGATGTCTCCTGGTGTGCTAACGTTTGCCA
TCATATGGCCTTTTATTGCACAGTGGTTGGTGTATTTATACTATCAAAGAAGGAGAAGAAAACTGAACTGAGGAAAATGG
TGTTTTCCTCAAGAAGAATACTGGAACTGGATGACCTCAGAATGAACTGGATTGTGGTGTTCACAAGAAAATCTTAGTTT
GTGATGATTACATTGCTTTTTGTTGTCCAGTAGTTTAGTTTGTGTACATATATACACATATATATTTTGCACTACACAAA
CGATAACATTTTAAGGACTAATATTGCTGATACTTGAATAATCAATCTCTACTAGGTTATAAGTAGTATACACAGATTTA
CCCTGCCCTTGAACTTGAAGGACATTAAATTATTAATGATCATTTGGTAACATGTTTACCTGATTATCTTCCATAGAGTA
ACATAAGCTGCTTTTCAAAGGTACCATTGTGATAATGAGATCAAATTTATAAGTTATTATTTTTAATTTTCTAAATTAAA
TAAAAGAAAGAATGCAAACCAGGAGTGAATTTCAAATGAGATGTAATCGACTTTATATCTTAGTCACGGAGTTGCCATGG
CATGTAGTAGAAAACCACAGGAAGAATGGTCATATTCACTTTGTGGGCTGCCCATAATCTTTCTTGGGCATTCACAACTC
TTGAGTTTGGTGTTCAGGCATCATTATTAAAAAGTGGAGTCCTATGTACCAGACTGAGTTTTTACAAATGATTTGCAGGC
TAGACATAACCCACTGATGGAAATGGTGAAGAATGAGCTTCATGTAGGTTTAAAAGTGTATTCTGAGCCTGTAGATGATT
AATCAGGTTTTTATTCAATCATATAAATGTATTCCTTTGTAATCATTTTTGTTTAACTGAGGATATCTAGTGTCTGCTTC
ATAGGGTGCTTTGAAATATAAAATGAAAACTTATTTATACTGTTTTTACAACAGTCAAAAAGGAAACACACGAAAATCAC
TTTTCTGCAACTGATGAACTATATAGACTGGACTCTTAACCTCTTAGTGCCTCAGTTTTTCCTTCGGGGTATACTATTTT
AGATACACCTACTTCACAGGAGTACTGGAGGGATTTGCAAGCTAACGGGCCAAATGCTTCCATAGGAATAAGGCATGCCC
AGCTAACAGAAATTTAAGTCCCTCTTCCCCACCTCTCTCATCTAGACCAAAAAGAAGACTACAATTTAGATCCTTGGAGA
CTTTTCCTGTATGCCTTCCAAAACTTCCTGCTCACTTAGATGTACCTTGTGCTTTGAAGATTTCTTTGATGCAGTCTTTG
TAACAAATAATTTTTTTTTTTAAATGGAGTTTCTCTCTTGTTGCCCAGGCTGGAGGGCAATGGCGTGATCTTGGCTCACT
GCAACTTCCACCTCCTGGGTTCAAGTGATTCTCCTGCCTCACTCCCAAGTAGCTGGGATTACAGGCATGCACCACCATGC
CTGGCTAATTTTGTATTTTTAGTAGAGACGGGGTTTCTCCATGTTGGTAAGGCTGATCTCAAACTCCCAACCTCAGGTGA
TCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTAGAGGCGTGAGCCACTGCACCTGGCCTGTAGCAAATAATTTTTAAG
CATTCTCAAGATGTATGATGTTGGGTTTAACATCATATGTTCACAGTGTTTTAAATAAGAAATAATCTGTCTTTAGTATA
CAGGATGGTTCTTTGTTTGCCTGAAAAGTATAAGAATACAATTTACTTTTCCCAAACTCTTTTTCCTTATTTTTTTCTTT
CAAATAACTTTAATTATTTAATCCCATACTGATTAAAATATGTCTGTCTAAAGGGATCTACTACTATTTGCTTTTAAAAA
AGTGTTCCCTATTTATTTAGGAAAAAGTGAAGCAAGCAACTGAATTTATATGTAAAAATAACATTTAGACCTGTGGATCA
AAGATTATTTTCAAAAATGAGATATCAATTTGGTATCAGTAAAAAGTGTCTTTTCTCTTAGTATATAAATAAATAAAATG
CAAATTAAGAGTGTAAGCAGTTGTGAATCATTACAAGCAGTCTGGGATAACTAACTCCTAGGGATAACCATTTCTATCCT
GTTCCAAAGTCTTATTTTATAGTTTGAAAAGCACTTTGCACACAGTTCTTGTATATAAAAGTAAAGATGTAATTATAGGA
TATAGTGTTCCTGCTTTGTTTAATAAGAACCTCATTTAAACTTGACAGCTATGGTATTTTTTTAATGATAACTTCTTTTC
TGTTTATTGTAAAACTAAGTTAAAAATAAAAGGTTAATGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 8)

The homology between NOV1a and its variants is shown graphically in the ClustalW analysis shown in Table 6. In the ClustalW alignment of NOV1a protein and its variants, the black outlined amino acid residues indicate regions of conserved sequence (*i.e.*, regions that may be required to preserve structural or functional properties), whereas non-highlighted amino acid residues are less conserved and can potentially be altered to a much broader extent without altering protein structure or function.

Table 6: Clustal W of NOV1 proteins



Based on the relatedness among NOV1, its variants, and endozepine-related protein precursor-like proteins, the disclosed NOV1 proteins are novel members of the endozepine-related protein precursor-like protein family. Therefore, the nucleic acids and proteins of the inventions are useful in potential therapeutic applications implicated in various pathologies and disorders described and other pathologies and disorders related to aberrant function or aberrant expression of these endozepine-related protein precursor-like proteins.

NOV2

NOV2 includes a novel endozepine-related protein disclosed below. A disclosed NOV2 nucleic acid of 1677 nucleotides (also referred to as CG51523-04) encoding a novel endozepine-related protein is shown in Table 7A. An open reading frame was identified beginning with an ATG initiation codon at nucleotides 35-37 and ending with a TGA codon at nucleotides 1637-1639. A putative untranslated region upstream from the initiation codon and downstream from the termination codon is underlined in Table 7A. The start and stop codons are in bold letters.

Table 7A. NOV2 nucleotide sequence (SEQ ID NO:9).

CCTTTTGGGGCATGTTGATCCGCGGCTGCGCTCC**ATG**TTCCAGTTTCATGCAGGCTCTTGGGAAAG
CTGGTGCTGCTGCTGCCTGATTCCCGCCGACAGACCTTGGGACCGGGGCCAACACTGGCAGCTGGAGATGGCG
GACACGAGATCCGTGCACGAGACTAGGTTTGAGGCGGCCGTGAAGGTGATCCAGAGTTTGCCGAAGAATGGTT
CATTCCAGCCAACAAATGAAATGATGCTTAAATTTTATAGCTTCTATAAGCAGGCAACTGAAGGACCCTGTAA
ACTTTCAAGGCCTGGATTTTGGGATCCTATTGGAAGATATAAATGGGATGCTTGGAGTTCACTGGGTGATATG
ACCAAAGAGGAAGCCATGATTGCATATGTTGAAGAAATGAAAAAGATTATTGAAACTATGCCAATGACTGAGA
AAGTTGAAGAATTGCTGCGTGTCATAGGTCCATTTTATGAAATTGTCGAGGACAAAAAGAGTGGCAGGAGTTC
TGATATAACCTCAGTCCGACTGGAGAAAATCTCTAAATGTTTAGAAGATCTTGGTAATGTTCTCACTTCTACT
CCAAACGCCAAAACCGTTAATGGTAAAGCTGAAAGCAGTGACAGTGGAGCCGAGTCTGAGGAAGAAGAGGCCC
AAGAAGAAGTGAAAGGAGCAGAACAAAGTGATAATGATAAGAAAATGATGAAGAAGTCAGCAGACCATAAGAA
TTTGGAAGTCATTGTCACTAATGGCTATGATAAAGATGGCTTTGTTCAGGATATACAGAATGACATTCATGCC
AGTTCTTCCCTGAATGGCAGAAGCACTGAAGAAGTAAAGCCCATTGATGAAAACTTGGGGCAAACTGGAAAAT
CTGCTGTTTGCATTCACCAAGATATAAATGATGATCATGTTGAAGATGTTACAGGAATTCAGCATTTGACAAG
CGATTCAGACAGTGAAGTTTACTGTGATTCTATGGAACAATTTGGACAAGAAGAGTCTTTAGACAGCTTTACG
TCCAACAATGGACCATTTCAGTATTACTTGGGTGGTCATTCCAGTCAACCCATGGAAAATTCTGGATTTCGTG
AAGATATTCAAGTACCTCCTGGAAATGGCAACATTGGGAATATGCAGGTGGTTGCAGTTGAAGGAAAAGGTGA
AGTCAAGCATGGAGGAGAAGATGGCAGGAATAACAGCGGAGCACTACACCGGGAGAAGCGAGGCGGAGAAACT
GACGAATTCTCTAATGTTAGAAGAGGAAGAGGACATAGGATGCAACACTTGAGCGAAGGAACCAAGGGCCGGC
AGGTGGGAAGTGGAGGTGATGGGGAGCGCTGGGGCTCCGACAGAGGGTCCCGAGGCAGCCTCAATGAGCAGAT
CGCCCTCGTGCTGATGAGACTGCAGGAGGACATGCAGAATGTCCTTCAGAGACTGCAGAAACTGGAAACGCTG
ACTGCTTTGCAGGCAAAATCATCAACATCAACATTGCAGACTGCTCCTCAGCCCACCTCACAGAGACCATCTT
GGTGGCCCTTCGAGATGTCTCCTGGTGTGCTAACGTTTGCCATCATATGGCCTTTTATTGCACAGTGGTTGGT
GTATTTATACTATCAAAGAAGGAGAAGAAAACTGAAC**TGA**GGAAAATGGTGTTTTCCTCAAGAAGAATACTGG
AACTG(SEQ ID NO: 9)

In all BLAST alignments herein, the "E-value" or "Expect" value is a numeric indication of the probability that the aligned sequences could have achieved their similarity to the BLAST query sequence by chance alone, within the database that was searched. For example, the probability that the subject ("Sbjct") retrieved from the NOV2 BLAST analysis, e.g., human CAC24877 mRNA, matched the Query NOV2 sequence purely by chance is 7.4e-293. The Expect value (E) is a parameter that describes the number of hits one can "expect" to see just by chance when searching a database of a particular size. It decreases exponentially with the Score (S) that is assigned to a match between two sequences. Essentially, the E value describes the random background noise that exists for matches between sequences.

The Expect value is used as a convenient way to create a significance threshold for reporting results. The default value used for blasting is typically set to 0.0001. In BLAST 2.0, the Expect value is also used instead of the P value (probability) to report the significance of matches. For example, an E value of one assigned to a hit can be interpreted as meaning that in a database of the current size one might expect to see one match with a similar score simply by chance. An E value of zero means that one would not expect to see any matches with a similar score simply by chance. See, e.g., http://www.ncbi.nlm.nih.gov/Education/BLASTinfo/. Occasionally, a string of X's or N's will result from a BLAST search. This is a result of automatic filtering of the query for low-complexity sequence that is performed to prevent artifactual hits. The filter substitutes any low-complexity sequence that it finds with the letter "N" in nucleotide sequence (e.g., "NNNNNNNNNNNN") or the letter "X" in protein sequences (e.g., "XXXXXXXXX"). Low-complexity regions can result in high scores that reflect compositional bias rather than significant position-by-position alignment. (Wootton and Federhen, Methods Enzymol 266:554-571, 1996).

The disclosed NOV2 polypeptide (SEQ ID NO:10) encoded by SEQ ID NO: 9 has 534 amino acid residues and is presented in Table 7B using the one-letter amino acid code. Signal P, Psort and Hydropathy results predict that NOV1 has a signal peptide and is likely to be localized in the plasma membrane with a certainty of 0.7300.

Exon linking data for NOV2 can be found below in Example 2. SNP analyses for NOV2 is described below in Example 3.

Table 7B. Encoded NOV2 protein sequence (SEQ ID NO:10).

```
MFQFHAGSWESWCCCCLIPADRPWDRGQHWQLEMADTRSVHETRFEAAVKVIQSLPKNGSFQPTNEMMLKFYSFYKQATEGPC
KLSRPGFWDPIGRYKWDAWSSLGDMTKEEAMIAYVEEMKKIIETMPMTEKVEELLRVIGPFYEIVEDKKSGRSSDITSVRLEK
ISKCLEDLGNVLTSTPNAKTVNGKAESSDSGAESEEEEAQEEVKGAEQSDNDKKMMKKSADHKNLEVIVTNGYDKDGFVQDIQ
NDIHASSSLNGRSTEEVKPIDENLGQTGKSAVCIHQDINDDHVEDVTGIQHLTSDSDSEVYCDSMEQFGQEESLDSFTSNNGP
FQYYLGGHSSQPMENSGFREDIQVPPGNGNIGNMQVVAVEGKGEVKHGGEDGRNNSGALHREKRGGETDEFSNVRRGRGHRMQ
```

HLSEGTKGRQVGSGGDGERWGSDRGSRGSLNEQIALVLMRLQEDMQNVLQRLQKLETLTALQAKSSTSTLQTAPQPTSQRPSW
WPFEMSPGVLTFAIIWPFIAQWLVYLYYQRRRRKLN(SEQ ID NO:10)

The full amino acid sequence of the disclosed NOV2 protein was found to have 530 of 534 amino acid residues (99%) identical to, and 531 of 534 amino acid residues (99%) similar to, the 534 amino acid residue ptnr:REMTRMBL-ACC:CAC24877 protein from sequence 23 of patent WO0078802. Public amino acid databases include the GenBank databases, SwissProt, PDB and PIR.

NOV2 maps to chromosome 10 and is expressed in at least the following tissues: : Brain, Colon, Foreskin, Kidney, Larynx, Lung, Mammary gland/Breast, Ovary, Pancreas, Placenta, Retina, Small Intestine, Spleen, Testis, Thalamus, and Uterus.

The amino acid sequence of NOV2 had high homology to other proteins as shown in Table 7C.

Table 7C. BLASTX results for NOV2

Sequences producing High-scoring Segment Pairs:	High Score	Smallest Sum Prob P(N)
patp:AAB48379 Human SEC12 protein sequence (clone ID 2093...	2815	5.3e-293
patp:AAU00399 Human secreted protein, POLY11 - Homo sapie...	2815	5.3e-293
patp:AAB48375 Human SEC8 protein sequence (clone ID 20936...	2809	2.3e-292
patp:AAB81816 Human endozepine-like ENDO6 SEQ ID NO: 23 -...	2769	4.0e-288
patp:AAB81816 Human endozepine-like ENDO6 SEQ ID NO: 23 -...	2769	4.0e-288

The disclosed NOV2 polypeptide also has homology to the amino acid sequences shown in the BLASTP data listed in Table 7D.

Table 7D. BLAST results for NOV2

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
CAC24877	Sequence 23 from Patent WO0078802/human	534	530/534 (99%)	531/534 (99%)	7.4e-293
CAC24873	Sequence 15 from Patent WO0078802/human	536	529/531 (99%)	529/531 (99%)	3.2e-292
P07106	Endozepine-related protein precursor/bovine	533	455/534 (85%),	484/534 (90%)	2.0e-251
Q9UFB5	Hypothetical 31.5kDa Protein/human	283	282/283 (99%)	282/283 (99%)	1.5e-152

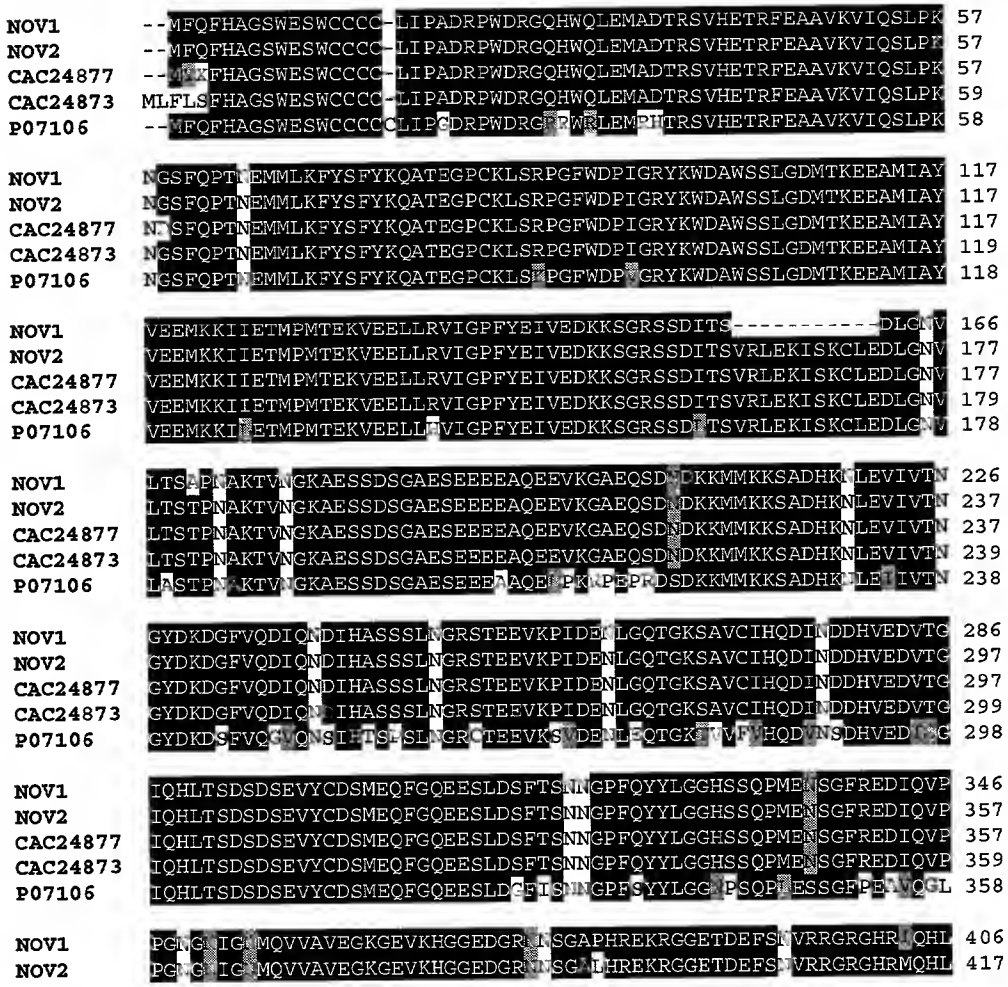
Q9CW41	1300014E15RIK Protein	504	283/405 (69%)	321/405 (79%)	8.2e-136

The homology between these and other sequences is shown graphically in the ClustalW analysis shown in Table 7E. In the ClustalW alignment of the NOV2 protein, as well as all other ClustalW analyses herein, the black outlined amino acid residues indicate regions of conserved sequence (*i.e.*, regions that may be required to preserve structural or functional properties), whereas non-highlighted amino acid residues are less conserved and can potentially be altered to a much broader extent without altering protein structure or function.

Table 7E. ClustalW Analyses of NOVX proteins

Information for the ClustalW proteins:

1. **NOV1** (SEQ ID NO: 2)
2. **NOV2** (SEQ ID NO: 10)
2. **CAC24877** (SEQ ID NO: 13)
3. **CAC24873** (SEQ ID NO: 14)
4. **P07106** (SEQ ID NO: 15)



Based on the relatedness among NOV1, NOV2 and endozepine-related protein precursor-like proteins, the disclosed NOVX proteins are novel members of the endozepine-related protein precursor-like protein family. Therefore, the nucleic acids and proteins of the inventions are useful in potential therapeutic applications implicated in various pathologies and disorders described and other pathologies and disorders related to aberrant function or aberrant expression of these endozepine-related protein precursor-like proteins.

The presence of identifiable domains in NOV2 was determined by searches using software algorithms such as PROSITE, DOMAIN, Blocks, Pfam, ProDomain, and Prints, and then determining the Interpro number by crossing the domain match (or numbers) using the Interpro website (http:www.ebi.ac.uk/ interpro). DOMAIN results for NOV2 as disclosed in Table 7F, were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST analyses. This BLAST analysis software samples domains found in theSmart and Pfam collections. For Table 7F and all successive DOMAIN sequence alignments, fully conserved single residues are indicated by black shading or by the sign (|) and "strong" semi-conserved residues are indicated by grey shading or by the sign (+). The "strong" group of conserved amino acid residues may be any one of the following groups of amino acids: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW.

Table 7F lists the domain description from DOMAIN analysis results against NOV2. This indicates that the NOV2 sequence has properties similar to those of other proteins known to contain this domain.

Table 7F. Domain Analysis of NOV2

```
ACBP (InterPro)  Acyl CoA binding protein

ACBP: domain 1 of 1, from 41 to 129: score 199.7, E = 4.4e-56

                *>lqedFeaAaekvkkLkknGpvkPSneekLkLYsLYKQATvGDvnter
                  ++++FeaA++++++L+knG+++P+ne++Lk+Ys+YKQAT+G+++++r
NOV2        41    HETRFEAAVKVIQSLPKNGSFQPTNEMMLKFYSFYKQATEGPCKLSR 87

                   PGmfDlkgrAKWDAWnelkGmSkeeAmkaYIakVeeLiakya<-* (SEQ ID NO: 12)
                   PG++D++gr+KWDAW++l++m+keeAm+aY+++++++i++++
NOV2        88     PGFWDPIGRYKWDAWSSLGDMTKEEAMIAYVEEMKKIIETMP    129
```

Daizepam binding inhibitor (DBI) is also known as acyl-CoA-binding protein (ACBP). The most highly conserved domain comprises 7 continguous residues that are identical in all known protein species from yeast, birds, and mammals. This domain constitutes the hydrophobic binding site for acyl-CoA-binding esters and is located within the second helical region of the molecule. The presence of a highly conservative gene in a primitive organism such as yeast supports its basic biologic role as an acyl-CoA-binding protein and suggests that many of the biologic functions attributed to it in higher organisms may result from its ability to interact with acyl-CoA.

Benzodiazepines modulate signal transduction at type A GABA (gamma-aminobutyric acid) receptors located in brain synapses. GABA is the predominant inhibitory neurotransmitter of the mammalian central nervous system. This receptor binds GABA, beta-carbolines, and benzodiazepines with high affinity and a chloride ion channel. Benzodiazepines prolong the chloride ion channel opening burst elicited by GABA and thereby enhance GABA-mediated inhibitory responses. This facilitation plays a role in reducing pathologic anxiety. An endogenous ligand has been identified that is recognized by the beta-carboline/benzodiazepine recognition site located in the GABA receptor. This ligand, diazepam binding inhibitor (DBI), is a protein of about 11 kD that displaces beta-carbolines and benzodiazepines bound to brain membrane fractions in vitro. DBI or a derivative small neuropeptide is thought to downregulate the effects of GABA.

Daizepam binding inhibitor/Acyl-CoA-binding protein(DBI/ACBP) is a small (10 Kd) protein that binds medium- and long-chain acyl-CoA esters with high affinity, and may act as an intra-cellular carrier of acyl-CoA esters. ACBP is a highly conserved protein that has been so far found in vertebrates, insects, plants and yeast. ACBP has a number of important physiological and biochemical functions: it is known as a diazepam binding inhibitor, as a putative neurotransmitter, as a regulator of insulin release from pancreatic cells, and as a mediator in corticotropin-dependent adrenal steroidogenesis. It is possible that the protein acts as a neuropeptide that takes part in the modulation of gamma-aminobutyric acid-ergic transmission. The structure of ACBP has been deduced by NMR spectroscopy and has been shown to be a mainly-alpha protein, consisting of 5 short alpha-helices and 3 connecting beta-strands.

Other proteins belonging to the ACBP family include mouse endozepine-like peptide (ELP); mammalian MA-DBI, a transmembrane protein of unknown function which has been found in mammals (like the one reported here in this invention); and human DRS-1, a protein of

unknown function that contains a N-terminal ACBP-like domain and a C-terminal enoyl-CoA isomerase/hydratase domain.

A polypeptide related to DBI, with similar binding activity to diazepam, has been isolated from human and bovine brain. This protein, called endozepine, contains 86 amino residues. Northern analysis using the cloned cDNA demonstrated that the message is expressed in heart, liver, and spleen, in addition to brain. Benzodiazepine receptors unassociated with the GABA receptor complex, and distinct from those seen in association with the central nervous system, have been identified in peripheral tissues.

Complementary DNAs (cDNAs) containing the entire coding sequence of endozepine, a putative ligand of the benzodiazepine receptor, were isolated from bovine and human cDNA libraries. These libraries were constructed using a novel oligonucleotide adapter molecule that allowed us to combine the use of G/C tailing with the preservation of the unique Eco RI site in the vector, lambda gt10. The amino acid sequences derived from these cDNA clones are identical to those previously determined for the purified proteins and are homologous to a related rat protein termed diazepam-binding inhibitor. The endozepine proteins are highly conserved, as illustrated by the finding that the nucleotide sequences of the coding regions are 93% conserved between the bovine and human forms. Analysis of these sequences indicates that endozepine is not, as expected, derived from a precursor molecule containing a transient signal peptide. Moreover, Northern analyses using the cloned cDNAs as hybridization probes indicate that the 650-nucleotide endozepine mRNA is expressed in a number of peripheral tissues in addition to brain. These observations may be consistent with a recent report describing the presence in peripheral tissues of benzodiazepine receptors on the outer mitochondrial membrane (Anholt et al., 1986). In addition to the endozepine cDNAs, a bovine cDNA clone was isolated that encodes a larger protein, a portion of which is homologous to endozepine. This related protein may be synthesized in a precursor form containing putative signal peptide and membrane-spanning domains.

Human endozepine, an 86 amino acid polypeptide, was originally isolated from human brain tissue as a putative ligand of the benzodiazepine receptor. Complete amino acid sequencing of the human and bovine proteins revealed significant homology with the partial sequence of diazepam binding inhibitor (DBI), a protein from rat brain. Both endozepine and DBI have been shown to elicit behavioral effects, suggesting that they function as pharmacologically-active ligands of the GABA (gamma-aminobutyric acid) receptor complex. Subsequent cDNA cloning of human and bovine endozepine, rat DBI and human DBI has

shown that these proteins are encoded by the same gene. A related cDNA, encoding a transmembrane protein of 533 amino acids with a domain homologous to DBI, has also been cloned from bovine brain.

Diazepam binding inhibitor (DBI) is a 10-kDa polypeptide that regulates mitochondrial steroidogenesis, glucose-induced insulin secretion, metabolism of acyl-CoA esters, and the action of gamma-aminobutyrate on GABA receptors. DBI is encoded by a muligene family of at least five members, but a single gene appears to account for the majority of DBI expression. DBI is expressed in a tissue-specific manner. Expression is found in central nervous system tissues and appears to extend to peripheral tissues rich in the peripheral type of high-affinity benzodiazepine recognition sites (Gray, PW, *et al.* 1986). To investigate the regulation of DBI gene expression, three positive clones were isolated from a rat genomic library. One of them contained a DBI genomic DNA fragment encompassing 4 kb of the 5' untranslated region, the first two exons, and part of the second intron of the DBI gene. Two other overlapping clones contained a processed DBI pseudogene. Several transcription initiation sites were detected by RNase protection and primer extension assays. Different tissues exhibited clear differences in the efficiencies of transcription startpoint usage. Transient expression experiments using DNA fragments of different length from the 5' untranslated region of the DBI gene showed that basal promoter activity required 146 bp of the proximal DBI sequence, whereas full activation was achieved with 423 bp of the 5' untranslated region. DNase I protection experiments with liver nuclear proteins demonstrated three protected regions at nt -387 to -333, -295 to -271, and -176 to -139 relative to the ATG initiation codon; in other tissues the pattern of protection was different. In gel shift assays the most proximal region (-176 to -139) was found to bind several general transcription factors as well as cell type-restricted nuclear proteins which may be related to specific regulatory patterns in different tissues. Thus, the DBI gene possesses some features of a housekeeping gene but also includes a variable regulation which appears to change with the function that it subserves in different cell types.

A trypsin-sensitive cholecystokinin-releasing peptide (CCK-RP) has been islolated from porcine and rat intestinal mucosa. The amino acid sequence of this peptide was determined to be identical to that of the diazepam-binding inhibitor (DBI). To test the role of DBI in pancreatic secretion and responses to feeding, pancreaticobiliary and intestinal cannula were used to divert bile-pancreatic juice from anesthetized rats. Within 2 hours, this treatment caused a 2-fold increase in pancreatic protein output and a >10-fold increase in plasma CCK. Luminal DBI levels increased 4-fold. At 5 hours after diversion of bile-pancreatic juice, each of these measures returned to basal levels. Intraduodenal infusion of peptone evoked a 5-fold increase in

the concentration of luminal DBI. In separate studies, it was demonstrated that intraduodenal administration of antiserum to a DBI peptide specifically abolished pancreatic secretion and the increase in plasma CCK levels after diversion of bile-pancreatic juice. To demonstrate that DBI mediates the postprandial rise in plasma CCK levels, it was shown that intraduodenal administration of 5% peptone induced dramatic increases in pancreatic secretion and plasma CCK, effects that could be blocked by intraduodenal administration of anti-DBI antiserum. Hence, DBI, a trypsin-sensitive CCK-RP secreted from the proximal small bowel, mediates the feedback regulation of pancreatic secretion and the postprandial release of CCK. Pancreatic proteases in the duodenum inhibit the release of CCK and thus exert feedback control of pancreatic exocrine secretion. Exclusion of proteases from the duodenum either by the diversion of bile-pancreatic juice or by the addition of protease inhibitors stimulates exocrine pancreatic secretion (Herzig, KH, *et al.* 1996).

Diazepam-binding inhibitor-derived peptides induce intracellular calcium changes and modulate human neutrophil function. The effects of two diazepam-binding inhibitor (DBI)-derived peptides, triakontatetraneuropeptide (DBI 17-50, TTN) and eiksoneuropeptide (DBI 51-70, ENP), on cytosolic free Ca2+ concentrations ([Ca2+]i), chemotaxis, superoxide anion (O2-) generation, and phagocytosis in human neutrophils were studied. Both TTN and ENP induced a rapid and transient rise of [Ca2+]i. The effect of TTN depended on the presence of extracellular Ca2+, whereas the effect of ENP also persisted after extracellular Ca2+ chelation. TTN induced neutrophil chemotaxis, stimulated O2- generation, and enhanced phagocytosis. ENP did not affect cell migration and oxidative metabolism but enhanced phagocytosis. Both peptides modulated N-formyl-methionyl-leucyl-phenylalanine- and phorbol myristate acetate-induced O2- generation. Because neutrophils express benzodiazepine receptors of the peripheral type (pBRs) and DBI-derived peptides may interact with such receptors, the possible role of pBRs in TTN- or ENP-induced effects were studied. The synthetic pBR ligand RO 5-4864 increased [Ca2+]i through extracellular Ca2+ influx and this effect was prevented by the pBR antagonist PK-11195. RO 5-4864, however, was ineffective on neutrophil migration and O2- generation and only slightly affected phagocytosis. Moreover, PK-11195 delayed the [Ca2+]i rise induced by TTN but did not significantly affect its extent, and had no effect on the [Ca2+]i rise induced by ENP. DBI-derived peptides induce [Ca2+]i changes and modulate neutrophil function mainly through pBR-independent pathways. In view of the wide cell and tissue distribution of DBI in the brain and in peripheral organs, modulation of neutrophil function by DBI-derived peptides may be relevant for both the neuroimmune network and the development and regulation of the inflammatory processes.

Peripheral benzodiazepine (BDZ) receptor (PBR) and diazepam-binding inhibitor/acyl-CoA-binding protein (DBI/ACBP) characterized as a ligand at central BDZ receptors, at PBR with involvement in the regulation of steroidogenesis, and as an intracellular acyl-CoA transporter, are both known to interact with BDZ in adult systems. Researchers investigated their expression after prenatal exposure to BDZ. Diazepam (1.25 mg/kg per day s.c.) was administered to time-pregnant Long Evans rats from gestational day (GD) 14 to 20. Expression of mRNAs encoding for PBR and for DBI/ACBP was studied in the same animals with (33)P-labeled 60 mer oligonucleotides (oligos) by in situ hybridization at GD20, and with (32)P-labeled oligos by Northern blot in steroidogenic and immune organs at postnatal day (PN) 14 and in adult offspring. Prenatal diazepam increased DBI/ACBP mRNA expression in male fetal adrenal and in fetal and PN14 testis. Thymus exhibited increased DBI/ACBP mRNA in male fetuses and in adult female offspring, and reduced organ weight at PN14 in both sexes. In female spleen, an increase in DBI/ACBP mRNA and a decrease in PBR mRNA was seen at PN14. Apart from the finding in spleen, no drug-induced changes in PBR mRNA were observed. The effects of prenatal diazepam were superimposed on treatment-independent sex differences in DBI/ACBP mRNA and PBR mRNA expression. Our data indicate that expression of DBI/ACBP mRNA in steroidogenic and immune organs can be affected by exposure to BDZ during ontogeny, while PBR mRNA expression appears to be less sensitive. They further reveal marked sex differences in the developmental patterns of the two proteins during pre- and postpubertal ontogeny.

The mechanisms underlying the increase in diazepam binding inhibitor (DBI) and its mRNA expression induced by nicotine (0.1 muM) exposure for 24 h were studied using mouse cerebral cortical neurons in primary culture. Nicotine-induced (0.1 muM) increases in DBI mRNA expression were abolished by hexamethonium, a nicotinic acetylcholine (nACh) receptor antagonist. Agents that stabilize the neuronal membrane, including tetrodotoxin (TTX), procainamide (a Na(+) channel inhibitor), and local anesthetics (dibucaine and lidocaine), dose-dependently inhibited the increased expression of DBI mRNA by nicotine. The nicotine-induced increase in DBI mRNA expression was inhibited by L-type voltage-dependent Ca(2+) channel (VDCC) inhibitors such as verapamil, calmodulin antagonist (W-7), and Ca(2+)/calmodulin-dependent protein kinase II (CAM II kinase) inhibitor (KN-62), whereas P/Q- and N-type VDCC inhibitors showed no effects. In addition, nicotine exposure for 24 h induced [3H]nicotine binding to the particulate fractions of the neurons with an increased B(max) value and no changes in K(d). Under these conditions, the 30 mM KCl- and nicotine-induced 45Ca(2+) influx into the nicotine-treated neurons was significantly higher than those

into non-treated neurons. These results suggest that the nicotine-stimulated increase in DBI mRNA expression is mediated by CAM II kinase activation resulting from the increase in intracellular Ca(2+) through L-type VDCCs subsequent to the neuronal membrane depolarization associated with nACh receptor activation.

Benzodiazepine receptors and DBI play a major role in regulating steroid production in both the adrenals and central nervous system, and may be involved in the activation of the hypothalamic-pituitary-adrenal axis in stress response. Preliminary findings regarding the presence of plasmatic benzodiazepine binding inhibitory activity (BBIA) in 14 psychiatric patients prompted us to investigate it further in larger samples of psychiatric patients (n = 44) and in healthy controls (n = 14). The results have shown that BBIA is present in all the subjects included with statistically significant differences between the patients and controls. The highest concentrations were found in the patients with no difference between anxious or depressed patients, and the lowest in the healthy controls. These findings indicate that BBIA might play a role in the pathophysiology of some manifestations of anxiety.

The neuropeptides diazepam binding inhibitor (DBI) and corticotropin-releasing hormone (CRH) elicit anxietylike symptoms when administered intracerebroventricularly to laboratory animals. Because of the similarities between the symptoms of certain anxiety states and the alcohol withdrawal syndrome, suggesting that increased secretion of either of these endogenous neuropeptides may, at least in part, be responsible for the symptoms of alcohol withdrawal. DBI and CRH concentrations were measured in cerebrospinal fluid (CSF) of 15 alcohol-dependent patients during acute withdrawal (Day 1) and again at 3 week's abstinence (Day 21). In addition, plasma concentrations of cortisol were measured to evaluate the relationship between pituitary-adrenal axis activation and CSF CRH concentrations. CSF CRH ($p < .04$), but not CSF DBI, was significantly higher on Day 1 than on Day 21. Although there was a significant decrease in plasma cortisol from Day 1 to Day 21 ($p < .001$), a significant correlation between CSF CRH and plasma cortisol concentrations was not observed at either time point. Neither CSF neuropeptide correlated with clinical measures of withdrawal severity. These tentative findings may implicate CRH, but not DBI, in the pathogenesis of alcohol withdrawal. Alternately, the central release of CRH and DBI may not be adequately reflected in lumbar CSF.

Because diazepam binding inhibitor (DBI) and its processing products coexist with gamma-aminobutyric acid (GABA) in several axon terminals, DBI immunoreactivity was measured in the cerebrospinal fluid (CSF) of individuals suffering from various neuropsychiatric disorders, that are believe to be associated with abnormalities of GABAergic

transmission. Increased amounts of DBI-like immunoreactivity were found in the CSF of patients suffering from severe depression with a severe anxiety component (Barbaccia, Costa, Ferrero, Guidotti, Roy, Sunderland, Pickar, Paul and Goodwin, 1986). Moreover, the amount of DBI and its processing products was found to be increased in the CSF of patients with hepatic encephalopathy (HE) (Rothstein, McKhann, Guarneri, Barbaccia, Guidotti and Costa, 1989; Guarneri, Berkovich, Guidotti and Costa, 1990). The clinical rating of HE correlated with the extent of the increase in DBI in CSF. Other lines of research suggest that DBI and DBI processing products may be important factors in behavioral adaptation to stress, acting via benzodiazepine (BZD) binding sites, located on mitochondria. DBI and its processing products, ODN and TTN, are present in high concentrations in the hypothalamus and in the amygdala, two areas of the brain that are important in regulating behavioral patterns associated with conflict situations, anxiety and stress. In CSF, the content of DBI changes in association with corticotropin releasing factor (CRF) (Roy, Pickar, Gold, Barbaccia, Guidotti, Costa and Linnoila, 1989). Finally DBI is preferentially concentrated in steroidogenic tissues and cells (adrenal cortical cells, Leydig cells of the testes and glial cells of the brain).

The disclosed NOVX nucleic acids of the invention encoding a endozepine-related protein precursor-like protein includes the nucleic acids whose sequences are provided herein, or a fragment thereof. The invention also includes a mutant or variant nucleic acid any of whose bases may be changed from the corresponding base shown herein while still encoding a protein that maintains its endozepine-related protein precursor-like activities and physiological functions, or a fragment of such a nucleic acid. The invention further includes nucleic acids whose sequences are complementary to those just described, including nucleic acid fragments that are complementary to any of the nucleic acids just described. The invention additionally includes nucleic acids or nucleic acid fragments, or complements thereto, whose structures include chemical modifications. Such modifications include, by way of nonlimiting example, modified bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized. These modifications are carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they may be used, for example, as antisense binding nucleic acids in therapeutic applications in a subject. In the mutant or variant nucleic acids, and their complements, up to about 10% percent of the bases may be so changed.

The disclosed NOVX proteins of the invention includes the endozepine-related protein precursor-like protein whose sequences are provided herein. The invention also includes mutant or variant proteins any of whose residues may be changed from the corresponding residues shown in SEQ ID NO: 2, 5, 7, or 10, while still encoding a protein that maintains its

endozepine-related protein precursor -like activities and physiological functions, or a functional fragment thereof. In the mutant or variant proteins, up to about 60% percent of the residues may be so changed.

The invention further encompasses antibodies and antibody fragments, such as F_{ab} or $(F_{ab})_2$, that bind immunospecifically to any of the proteins of the invention.

The above defined information for this invention suggests that thess NOVX endozepine-related protein precursor-like proteins may function as a member of a "endozepine-related protein precursor family". Therefore, the NOVX nucleic acids and proteins identified here may be useful in potential therapeutic applications implicated in (but not limited to) various pathologies and disorders as indicated below. The potential therapeutic applications for this invention include, but are not limited to: protein therapeutic, small molecule drug target, antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), diagnostic and prognostic marker, gene therapy (gene delivery/gene ablation), research tools, tissue regeneration in vivo and in vitro of all tissues and cell types composing (but not limited to) those defined here.

The NOVX nucleic acids and proteins of the invention are useful in potential therapeutic applications implicated in cancer including but not limited to various pathologies and disorders as indicated below. For example, a cDNA encoding the NOVX endozepine-related protein precursor-like proteins may be useful in gene therapy, and the NOVX endozepine-related protein precursor-like protein may be useful when administered to a subject in need thereof. By way of nonlimiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, stroke, tuberous sclerosis, hypercalceimia, Parkinson's disease, Huntington's disease, cerebral palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia-telangiectasia, leukodystrophies, behavioral disorders, addiction, anxiety, pain, neuroprotection, anxiety, depression, stress, immune dysfunction, alcoholism, obesity and diabetes. The NOVX nucleic acids encoding the endozepine-related protein precursor-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed.

NOVX nucleic acids and polypeptides are further useful in the generation of antibodies that bind immuno-specifically to the novel NOVX substances for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-NOVX Antibodies" section below. The disclosed NOVX proteins have multiple hydrophilic regions, each of which

can be used as an immunogen. In one embodiment, a contemplated NOVX epitope is from about amino acids 480 to 500. These novel proteins can be used in assay systems for functional analysis of various human disorders, which will help in understanding of pathology of the disease and development of new drug targets for various disorders.

NOVX Nucleic Acids and Polypeptides

One aspect of the invention pertains to isolated nucleic acid molecules that encode NOVX polypeptides or biologically active portions thereof. Also included in the invention are nucleic acid fragments sufficient for use as hybridization probes to identify NOVX-encoding nucleic acids (*e.g.*, NOVX mRNAs) and fragments for use as PCR primers for the amplification and mutation of NOVX nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule may be single-stranded or double-stranded, but preferably is comprised double-stranded DNA.

An NOVX nucleic acid can encode a mature NOVX polypeptide. As used herein, a "mature" form of a polypeptide or protein disclosed in the present invention is the product of a naturally occurring polypeptide or precursor form or proprotein. The naturally occurring polypeptide, precursor or proprotein includes, by way of nonlimiting example, the full-length gene product, encoded by the corresponding gene. Alternatively, it may be defined as the polypeptide, precursor or proprotein encoded by an ORF described herein. The product "mature" form arises, again by way of nonlimiting example, as a result of one or more naturally occurring processing steps as they may take place within the cell, or host cell, in which the gene product arises. Examples of such processing steps leading to a "mature" form of a polypeptide or protein include the cleavage of the N-terminal methionine residue encoded by the initiation codon of an ORF, or the proteolytic cleavage of a signal peptide or leader sequence. Thus a mature form arising from a precursor polypeptide or protein that has residues 1 to N, where residue 1 is the N-terminal methionine, would have residues 2 through N remaining after removal of the N-terminal methionine. Alternatively, a mature form arising from a precursor polypeptide or protein having residues 1 to N, in which an N-terminal signal sequence from residue 1 to residue M is cleaved, would have the residues from residue M+1 to residue N remaining. Further as used herein, a "mature" form of a polypeptide or protein may arise from a step of post-translational modification other than a proteolytic cleavage event. Such additional processes include, by way of non-limiting example, glycosylation, myristoylation or

phosphorylation. In general, a mature polypeptide or protein may result from the operation of only one of these processes, or a combination of any of them.

The term "probes", as utilized herein, refers to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt), 100 nt, or as many as approximately, *e.g.*, 6,000 nt, depending upon the specific use. Probes are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are generally obtained from a natural or recombinant source, are highly specific, and much slower to hybridize than shorter-length oligomer probes. Probes may be single- or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

The term "isolated" nucleic acid molecule, as utilized herein, is one, which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5'- and 3'-termini of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated NOVX nucleic acid molecules can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell/tissue from which the nucleic acid is derived (*e.g.*, brain, heart, liver, spleen, etc.). Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence SEQ ID NO: 1, 3, 4, 6, 8, or 9, or a complement of this aforementioned nucleotide sequence, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of SEQ ID NO: 1, 3, 4, 6, 8, or 9 as a hybridization probe, NOVX molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, *et al.*, (eds.), MOLECULAR CLONING: A LABORATORY MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993.)

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides

corresponding to NOVX nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues, which oligonucleotide has a sufficient number of nucleotide bases to be used in a PCR reaction. A short oligonucleotide sequence may be based on, or designed from, a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a nucleic acid sequence having about 10 nt, 50 nt, or 100 nt in length, preferably about 15 nt to 30 nt in length. In one embodiment of the invention, an oligonucleotide comprising a nucleic acid molecule less than 100 nt in length would further comprise at least 6 contiguous nucleotides SEQ ID NO: 1, 3, 4, 6, 8, or 9, or a complement thereof. Oligonucleotides may be chemically synthesized and may also be used as probes.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in SEQ ID NO: 1, 3, 4, 6, 8, or 9, or a portion of this nucleotide sequence (*e.g.*, a fragment that can be used as a probe or primer or a fragment encoding a biologically-active portion of an NOVX polypeptide). A nucleic acid molecule that is complementary to the nucleotide sequence shown SEQ ID NO: 1, 3, 4, 6, 8, or 9 is one that is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO: 1, 3, 4, 6, 8, or 9 that it can hydrogen bond with little or no mismatches to the nucleotide sequence shown SEQ ID NO: 1, 3, 4, 6, 8, or 9 thereby forming a stable duplex.

As used herein, the term "complementary" refers to Watson-Crick or Hoogsteen base pairing between nucleotides units of a nucleic acid molecule, and the term "binding" means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, van der Waals, hydrophobic interactions, and the like. A physical interaction can be either direct or indirect. Indirect interactions may be through or due to the effects of another polypeptide or compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

Fragments provided herein are defined as sequences of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific hybridization in the case of nucleic acids or for specific recognition of an epitope in the case of amino acids, respectively, and are at most some portion less than a full length sequence. Fragments may be derived from any contiguous portion of a nucleic acid or amino acid

sequence of choice. Derivatives are nucleic acid sequences or amino acid sequences formed from the native compounds either directly or by modification or partial substitution. Analogs are nucleic acid sequences or amino acid sequences that have a structure similar to, but not identical to, the native compound but differs from it in respect to certain components or side chains. Analogs may be synthetic or from a different evolutionary origin and may have a similar or opposite metabolic activity compared to wild type. Homologs are nucleic acid sequences or amino acid sequences of a particular gene that are derived from different species.

Derivatives and analogs may be full length or other than full length, if the derivative or analog contains a modified nucleic acid or amino acid, as described below. Derivatives or analogs of the nucleic acids or proteins of the invention include, but are not limited to, molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the invention, in various embodiments, by at least about 70%, 80%, or 95% identity (with a preferred identity of 80-95%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent conditions. *See e.g.* Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below.

A "homologous nucleic acid sequence" or "homologous amino acid sequence," or variations thereof, refer to sequences characterized by a homology at the nucleotide level or amino acid level as discussed above. Homologous nucleotide sequences encode those sequences coding for isoforms of NOVX polypeptides. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the invention, homologous nucleotide sequences include nucleotide sequences encoding for an NOVX polypeptide of species other than humans, including, but not limited to: vertebrates, and thus can include, *e.g.*, frog, mouse, rat, rabbit, dog, cat cow, horse, and other organisms. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the exact nucleotide sequence encoding human NOVX protein. Homologous nucleic acid sequences include those nucleic acid sequences that encode conservative amino acid substitutions (see below) in SEQ ID NO: 1, 3, 4, 6, 8, or 9, as well as a polypeptide possessing NOVX biological activity. Various biological activities of the NOVX proteins are described below.

An NOVX polypeptide is encoded by the open reading frame ("ORF") of an NOVX nucleic acid. An ORF corresponds to a nucleotide sequence that could potentially be translated into a polypeptide. A stretch of nucleic acids comprising an ORF is uninterrupted by a stop codon. An ORF that represents the coding sequence for a full protein begins with an ATG "start" codon and terminates with one of the three "stop" codons, namely, TAA, TAG, or TGA. For the purposes of this invention, an ORF may be any part of a coding sequence, with or without a start codon, a stop codon, or both. For an ORF to be considered as a good candidate for coding for a *bona fide* cellular protein, a minimum size requirement is often set, *e.g.*, a stretch of DNA that would encode a protein of 50 amino acids or more.

The nucleotide sequences determined from the cloning of the human NOVX genes allows for the generation of probes and primers designed for use in identifying and cloning NOVX homologues in other cell types, *e.g.* from other tissues, as well as NOVX homologues from other vertebrates. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 consecutive sense strand nucleotide sequence SEQ ID NO: 1, 3, 4, 6, 8, or 9; or an anti-sense strand nucleotide sequence of SEQ ID NO: 1, 3, 4, 6, 8, or 9; or of a naturally occurring mutant of SEQ ID NO: 1, 3, 4, 6, 8, or 9.

Probes based on the human NOVX nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In various embodiments, the probe further comprises a label group attached thereto, *e.g.* the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissues which mis-express an NOVX protein, such as by measuring a level of an NOVX-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting NOVX mRNA levels or determining whether a genomic NOVX gene has been mutated or deleted.

"A polypeptide having a biologically-active portion of an NOVX polypeptide" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. A nucleic acid fragment encoding a "biologically-active portion of NOVX" can be prepared by isolating a portion SEQ ID NO: 1, 3, 4, 6, 8, or 9, that encodes a polypeptide having an NOVX biological activity (the biological activities of the NOVX proteins are described below), expressing the encoded portion of NOVX protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of NOVX.

NOVX Nucleic Acid and Polypeptide Variants

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequences shown in SEQ ID NO: 1, 3, 4, 6, 8, or 9 due to degeneracy of the genetic code and thus encode the same NOVX proteins as that encoded by the nucleotide sequences shown in SEQ ID NO: 1, 3, 4, 6, 8, or 9. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO: 2, 5, 7, or 10.

In addition to the human NOVX nucleotide sequences shown in SEQ ID NO: 1, 3, 4, 6, 8, or 9, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the NOVX polypeptides may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the NOVX genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame (ORF) encoding an NOVX protein, preferably a vertebrate NOVX protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the NOVX genes. Any and all such nucleotide variations and resulting amino acid polymorphisms in the NOVX polypeptides, which are the result of natural allelic variation and that do not alter the functional activity of the NOVX polypeptides, are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding NOVX proteins from other species, and thus that have a nucleotide sequence that differs from the human SEQ ID NO: 1, 3, 4, 6, 8, or 9 are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the NOVX cDNAs of the invention can be isolated based on their homology to the human NOVX nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 6 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1, 3, 4, 6, 8, or 9. In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250, 500, 750, 1000, 1500, or 2000 or more nucleotides in length. In yet another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which

nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other.

Homologs (*i.e.*, nucleic acids encoding NOVX proteins derived from species other than human) or other related sequences (*e.g.*, paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5 °C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at Tm, 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (*e.g.*, 10 nt to 50 nt) and at least about 60°C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in Ausubel, *et al.*, (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain hybridized to each other. A non-limiting example of stringent hybridization conditions are hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C, followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequences SEQ ID NO: 1, 3, 4, 6, 8, or 9, corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In a second embodiment, a nucleic acid sequence that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1, 3, 4, 6, 8, or 9, or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 mg/ml denatured salmon sperm DNA at 55°C, followed by one or more washes in 1X SSC, 0.1% SDS at 37°C. Other conditions of moderate stringency that may be used are well-known within the art. *See, e.g.*, Ausubel, et *al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990; GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that is hybridizable to the nucleic acid molecule comprising the nucleotide sequences SEQ ID NO: 1, 3, 4, 6, 8, or 9, or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10% (wt/vol) dextran sulfate at 40°C, followed by one or more washes in 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at 50°C. Other conditions of low stringency that may be used are well known in the art (*e.g.*, as employed for cross-species hybridizations). *See, e.g.,* Ausubel, *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo and Weinberg, 1981. *Proc Natl Acad Sci USA* **78**: 6789-6792.

Conservative Mutations

In addition to naturally-occurring allelic variants of NOVX sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences SEQ ID NO: 1, 3, 4, 6, 8, or 9, thereby leading to changes in the amino acid sequences of the encoded NOVX proteins, without altering the functional ability of said NOVX proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence SEQ ID NO: 2, 5, 7, or 10. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequences of the NOVX proteins without altering their biological activity, whereas an "essential" amino acid residue is required for such biological activity. For example, amino acid residues that are conserved among the NOVX proteins of the invention are

predicted to be particularly non-amenable to alteration. Amino acids for which conservative substitutions can be made are well-known within the art.

Another aspect of the invention pertains to nucleic acid molecules encoding NOVX proteins that contain changes in amino acid residues that are not essential for activity. Such NOVX proteins differ in amino acid sequence from SEQ ID NO: 1, 3, 4, 6, 8, or 9, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 45% homologous to the amino acid sequences SEQ ID NO: 2, 5, 7, or 10. Preferably, the protein encoded by the nucleic acid molecule is at least about 60% homologous to SEQ ID NO: 2, 5, 7, or 10; more preferably at least about 70% homologous SEQ ID NO: 2, 5, 7, or 10; still more preferably at least about 80% homologous to SEQ ID NO: 2, 5, 7, or 10; even more preferably at least about 90% homologous to SEQ ID NO: 2, 5, 7, or 10; and most preferably at least about 95% homologous to SEQ ID NO: 2, 4, 7, or 10.

An isolated nucleic acid molecule encoding an NOVX protein homologous to the protein of SEQ ID NO: 2, 4, 7, or 10 can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO: 1, 3, 4, 6, 8, or 9, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into SEQ ID NO: 1, 3, 4, 6, 8, or 9 by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted, non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined within the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted non-essential amino acid residue in the NOVX protein is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of an NOVX coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for NOVX biological activity to identify mutants that retain activity. Following mutagenesis SEQ

ID NO: 1, 3, 4, 6, 8, or 9, the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

The relatedness of amino acid families may also be determined based on side chain interactions. Substituted amino acids may be fully conserved "strong" residues or fully conserved "weak" residues. The "strong" group of conserved amino acid residues may be any one of the following groups: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW, wherein the single letter amino acid codes are grouped by those amino acids that may be substituted for each other. Likewise, the "weak" group of conserved residues may be any one of the following: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, VLIM, HFY, wherein the letters within each group represent the single letter amino acid code.

In one embodiment, a mutant NOVX protein can be assayed for (*i*) the ability to form protein:protein interactions with other NOVX proteins, other cell-surface proteins, or biologically-active portions thereof, (*ii*) complex formation between a mutant NOVX protein and an NOVX ligand; or (*iii*) the ability of a mutant NOVX protein to bind to an intracellular target protein or biologically-active portion thereof; (*e.g.* avidin proteins).

In yet another embodiment, a mutant NOVX protein can be assayed for the ability to regulate a specific biological function (*e.g.*, regulation of insulin release).

Antisense Nucleic Acids

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1, 3, 4, 6, 8, or 9, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein (*e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence). In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire NOVX coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of an NOVX protein of SEQ ID NO: 2, 5, 7, or 10, or antisense nucleic acids complementary to an NOVX nucleic acid sequence of SEQ ID NO: 1, 3, 4, 6, 8, or 9, are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding an NOVX protein. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding the

NOVX protein. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding the NOVX protein disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of NOVX mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of NOVX mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of NOVX mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally-occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids (*e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used).

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and genomic DNA encoding an NOVX protein to thereby inhibit expression of the protein (*e.g.*, by inhibiting transcription and translation). The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface (*e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens). The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient nucleic acid molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other. *See, e.g.,* Gaultier, *et al.*, 1987. *Nucl. Acids Res.* **15**: 6625-6641. The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (*See, e.g.,* Inoue, *et al.* 1987. *Nucl. Acids Res.* **15**: 6131-6148) or a chimeric RNA-DNA analogue (*See, e.g.,* Inoue, *et al.*, 1987. *FEBS Lett.* **215**: 327-330.

Ribozymes and PNA Moieties

Nucleic acid modifications include, by way of non-limiting example, modified bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized. These modifications are carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they may be used, for example, as antisense binding nucleic acids in therapeutic applications in a subject.

In one embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes as described in Haselhoff and Gerlach 1988. *Nature* 334: 585-591) can be used to catalytically cleave NOVX mRNA transcripts to thereby

inhibit translation of NOVX mRNA. A ribozyme having specificity for an NOVX-encoding nucleic acid can be designed based upon the nucleotide sequence of an NOVX cDNA disclosed herein (*i.e.*, SEQ ID NO: 1, 3, 4, 6, 8, or 9). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an NOVX-encoding mRNA. *See*, *e.g.*, U.S. Patent 4,987,071 to Cech, *et al.* and U.S. Patent 5,116,742 to Cech, *et al.* NOVX mRNA can also be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. *See, e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, NOVX gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the NOVX nucleic acid (*e.g.*, the NOVX promoter and enhancers) to form triple helical structures that prevent transcription of the NOVX gene in target cells. *See, e.g.*, Helene, 1991. *Anticancer Drug Des.* 6: 569-84; Helene, *et al.* 1992. *Ann. N.Y. Acad. Sci.* 660: 27-36; Maher, 1992. *Bioassays* 14: 807-15.

In various embodiments, the NOVX nucleic acids can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids. *See, e.g.*, Hyrup, *et al.*, 1996. *Bioorg Med Chem* 4: 5-23. As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics (*e.g.*, DNA mimics) in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup, *et al.*, 1996. *supra*; Perry-O'Keefe, *et al.*, 1996. *Proc. Natl. Acad. Sci. USA* 93: 14670-14675.

PNAs of NOVX can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of NOVX can also be used, for example, in the analysis of single base pair mutations in a gene (*e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S_1 nucleases (*See,* Hyrup, *et al.*, 1996.*supra*); or as probes or primers for DNA sequence and hybridization (*See,* Hyrup, *et al.*, 1996, *supra*; Perry-O'Keefe, *et al.*, 1996. *supra*).

In another embodiment, PNAs of NOVX can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of

PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of NOVX can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes (*e.g.*, RNase H and DNA polymerases) to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (*see,* Hyrup, et al., 1996. *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup, *et al.,* 1996. *supra* and Finn, *et al.*, 1996. *Nucl Acids Res* 24: 3357-3363. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA. *See, e.g.,* Mag, *et al.*, 1989. *Nucl Acid Res* 17: 5973-5988. PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment. *See, e.g.,* Finn, *et al.*, 1996. *supra.* Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. *See, e.g.,* Petersen, *et al.*, 1975. *Bioorg. Med. Chem. Lett.* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (*see, e.g.*, Letsinger, *et al.*, 1989. *Proc. Natl. Acad. Sci. U.S.A.* 86: 6553-6556; Lemaitre, *et al.*, 1987. *Proc. Natl. Acad. Sci.* 84: 648-652; PCT Publication No. WO88/09810) or the blood-brain barrier (*see, e.g.*, PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (*see, e.g.,* Krol, *et al.*, 1988. *BioTechniques* 6:958-976) or intercalating agents (*see, e.g.*, Zon, 1988. *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, and the like.

NOVX Polypeptides

A polypeptide according to the invention includes a polypeptide including the amino acid sequence of NOVX polypeptides whose sequences are provided in SEQ ID NO: 2, 5, 7, or 10. The invention also includes a mutant or variant protein any of whose residues may be changed from the corresponding residues shown in SEQ ID NO: 2, 5, 7, or 10 while still encoding a protein that maintains its NOVX activities and physiological functions, or a functional fragment thereof.

In general, an NOVX variant that preserves NOVX-like function includes any variant in which residues at a particular position in the sequence have been substituted by other amino acids, and further include the possibility of inserting an additional residue or residues between two residues of the parent protein as well as the possibility of deleting one or more residues from the parent sequence. Any amino acid substitution, insertion, or deletion is encompassed by the invention. In favorable circumstances, the substitution is a conservative substitution as defined above.

One aspect of the invention pertains to isolated NOVX proteins, and biologically-active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-NOVX antibodies. In one embodiment, native NOVX proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, NOVX proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, an NOVX protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" polypeptide or protein or biologically-active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the NOVX protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of NOVX proteins in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly-produced. In one embodiment, the language "substantially free of cellular material" includes preparations of NOVX proteins having less than about 30% (by dry weight) of non-NOVX proteins (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-NOVX proteins, still more preferably less than about 10% of non-NOVX proteins, and most preferably less than about 5% of non-NOVX proteins. When the NOVX protein or biologically-active portion thereof is recombinantly-produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the NOVX protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of NOVX proteins in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of

NOVX proteins having less than about 30% (by dry weight) of chemical precursors or non-NOVX chemicals, more preferably less than about 20% chemical precursors or non-NOVX chemicals, still more preferably less than about 10% chemical precursors or non-NOVX chemicals, and most preferably less than about 5% chemical precursors or non-NOVX chemicals.

Biologically-active portions of NOVX proteins include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequences of the NOVX proteins (*e.g.*, the amino acid sequence shown in SEQ ID NO: 2, 5, 7, or 10) that include fewer amino acids than the full-length NOVX proteins, and exhibit at least one activity of an NOVX protein. Typically, biologically-active portions comprise a domain or motif with at least one activity of the NOVX protein. A biologically-active portion of an NOVX protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acid residues in length.

Moreover, other biologically-active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native NOVX protein.

In an embodiment, the NOVX protein has an amino acid sequence shown SEQ ID NO: 2, 5, 7, or 10. In other embodiments, the NOVX protein is substantially homologous to SEQ ID NO: 2, 5, 7, or 10, and retains the functional activity of the protein of SEQ ID NO: 2, 5, 7, or 10, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail, below. Accordingly, in another embodiment, the NOVX protein is a protein that comprises an amino acid sequence at least about 45% homologous to the amino acid sequence SEQ ID NO: 2, 5, 7, or 10, and retains the functional activity of the NOVX proteins of SEQ ID NO: 2, 5, 7, or 10.

Determining Homology Between Two or More Sequences

To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

The nucleic acid sequence homology may be determined as the degree of identity between two sequences. The homology may be determined using computer programs known in the art, such as GAP software provided in the GCG program package. *See,* Needleman and Wunsch, 1970. *J Mol Biol* 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NO: 1, 3, 4, 6, 8, or 9. The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.,* A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.,* the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region.

Chimeric and Fusion Proteins

The invention also provides NOVX chimeric or fusion proteins. As used herein, an NOVX "chimeric protein" or "fusion protein" comprises an NOVX polypeptide operatively-linked to a non-NOVX polypeptide. An "NOVX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to an NOVX protein SEQ ID NO: 2, 5, 7, or 10, whereas a "non-NOVX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the NOVX protein, *e.g.*, a protein that is different from the NOVX protein and that is derived from the same or a different organism. Within an NOVX fusion protein the NOVX polypeptide can correspond to all or a portion of an NOVX protein. In one embodiment, an NOVX fusion protein comprises at least one biologically-active portion of an NOVX protein. In another embodiment, an NOVX fusion protein comprises at least two biologically-active portions of an NOVX protein. In yet another embodiment, an NOVX fusion protein comprises at least three biologically-active portions of

an NOVX protein. Within the fusion protein, the term "operatively-linked" is intended to indicate that the NOVX polypeptide and the non-NOVX polypeptide are fused in-frame with one another. The non-NOVX polypeptide can be fused to the N-terminus or C-terminus of the NOVX polypeptide.

In one embodiment, the fusion protein is a GST-NOVX fusion protein in which the NOVX sequences are fused to the C-terminus of the GST (glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant NOVX polypeptides.

In another embodiment, the fusion protein is an NOVX protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and secretion of NOVX can be increased through use of a heterologous signal sequence.

In yet another embodiment, the fusion protein is an NOVX-immunoglobulin fusion protein in which the NOVX sequences are fused to sequences derived from a member of the immunoglobulin protein family. The NOVX-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between an NOVX ligand and an NOVX protein on the surface of a cell, to thereby suppress NOVX-mediated signal transduction *in vivo*. The NOVX-immunoglobulin fusion proteins can be used to affect the bioavailability of an NOVX cognate ligand. Inhibition of the NOVX ligand/NOVX interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (*e.g.* promoting or inhibiting) cell survival. Moreover, the NOVX-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-NOVX antibodies in a subject, to purify NOVX ligands, and in screening assays to identify molecules that inhibit the interaction of NOVX with an NOVX ligand.

An NOVX chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a

chimeric gene sequence (*see, e.g.,* Ausubel, *et al.* (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). An NOVX-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the NOVX protein.

NOVX Agonists and Antagonists

The invention also pertains to variants of the NOVX proteins that function as either NOVX agonists (*i.e.,* mimetics) or as NOVX antagonists. Variants of the NOVX protein can be generated by mutagenesis (*e.g.*, discrete point mutation or truncation of the NOVX protein). An agonist of the NOVX protein can retain substantially the same, or a subset of, the biological activities of the naturally occurring form of the NOVX protein. An antagonist of the NOVX protein can inhibit one or more of the activities of the naturally occurring form of the NOVX protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the NOVX protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the NOVX proteins.

Variants of the NOVX proteins that function as either NOVX agonists (*i.e.,* mimetics) or as NOVX antagonists can be identified by screening combinatorial libraries of mutants (*e.g.*, truncation mutants) of the NOVX proteins for NOVX protein agonist or antagonist activity. In one embodiment, a variegated library of NOVX variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of NOVX variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential NOVX sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of NOVX sequences therein. There are a variety of methods which can be used to produce libraries of potential NOVX variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential NOVX sequences. Methods for synthesizing degenerate oligonucleotides are well-known within the art. *See, e.g.,*

Narang, 1983. *Tetrahedron* 39: 3; Itakura, *et al.*, 1984. *Annu. Rev. Biochem.* 53: 323; Itakura, *et al.*, 1984. *Science* 198: 1056; Ike, *et al.*, 1983. *Nucl. Acids Res.* 11: 477.

Polypeptide Libraries

In addition, libraries of fragments of the NOVX protein coding sequences can be used to generate a variegated population of NOVX fragments for screening and subsequent selection of variants of an NOVX protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of an NOVX coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double-stranded DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S_1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, expression libraries can be derived which encodes N-terminal and internal fragments of various sizes of the NOVX proteins.

Various techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of NOVX proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify NOVX variants. *See, e.g.,* Arkin and Yourvan, 1992. *Proc. Natl. Acad. Sci. USA* 89: 7811-7815; Delgrave, *et al.*, 1993. *Protein Engineering* 6:327-331.

Anti-NOVX Antibodies

Also included in the invention are antibodies to NOVX proteins, or fragments of NOVX proteins. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab}, $F_{ab'}$ and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from

humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG_1, IgG_2, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated NOVX-related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of NOVX-related protein that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human NOVX-related protein sequence will indicate which regions of a NOVX-related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow and Lane, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

Polyclonal Antibodies

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

Monoclonal Antibodies

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular

species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, MONOCLONAL ANTIBODIES: PRINCIPLES AND PRACTICE, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, *J. Immunol.*, 133:3001 (1984); Brodeur et al., MONOCLONAL ANTIBODY PRODUCTION TECHNIQUES AND APPLICATIONS, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, *Anal. Biochem.*, 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, *Nature* 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

Humanized Antibodies

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-327 (1988); Verhoeyen et al., *Science*, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)).

Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein. Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by

transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991); Marks et al., *J. Mol. Biol.*, 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (*Bio/Technology* 10, 779-783 (1992)); Lonberg et al. (*Nature* 368 856-859 (1994)); Morrison (*Nature* 368, 812-13 (1994)); Fishwild et al,(*Nature Biotechnology* 14, 845-51 (1996)); Neuberger (*Nature Biotechnology* 14, 826 (1996)); and Lonberg and Huszar (*Intern. Rev. Immunol.* 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse™ as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent

No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

F_{ab} Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, *Nature*, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., *Methods in Enzymology*, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. $F(ab')_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., *Science* 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments

generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., *J. Exp. Med.* 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., *J. Immunol.* 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., *Proc. Natl. Acad. Sci. USA* 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., *J. Immunol.* 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., *J. Immunol.* 147:60 (1991).

Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to

focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a specific embodiment, selection of antibodies that are specific to a particular domain of an NOVX protein is facilitated by generation of hybridomas that bind to the fragment of an NOVX protein possessing such a domain. Thus, antibodies that are specific for a desired domain within an NOVX protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

Anti-NOVX antibodies may be used in methods known within the art relating to the localization and quantitation of an NOVX protein (*e.g.,* for use in measuring levels of the

NOVX protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for NOVX proteins, or derivatives, fragments, analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds (hereinafter "Therapeutics").

An anti-NOVX antibody (*e.g.*, monoclonal antibody) can be used to isolate an NOVX polypeptide by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-NOVX antibody can facilitate the purification of natural NOVX polypeptide from cells and of recombinantly-produced NOVX polypeptide expressed in host cells. Moreover, an anti-NOVX antibody can be used to detect NOVX protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the NOVX protein. Anti-NOVX antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

NOVX Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding an NOVX protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome.

Moreover, certain vectors are capable of directing the expression of genes to which they are operatively-linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively-linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably-linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell).

The term "regulatory sequence" is intended to includes promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, NOVX proteins, mutant forms of NOVX proteins, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of NOVX proteins in prokaryotic or eukaryotic cells. For example, NOVX proteins can be expressed in bacterial cells such as *Escherichia coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego,

Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *Escherichia coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (*i*) to increase expression of recombinant protein; (*ii*) to increase the solubility of the recombinant protein; and (*iii*) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson, 1988. *Gene* 67: 31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. *See, e.g.,* Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (*see, e.g.,* Wada, *et al.*, 1992. *Nucl. Acids Res.* 20: 2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the NOVX expression vector is a yeast expression vector. Examples of vectors for expression in yeast *Saccharomyces cerivisae* include pYepSec1 (Baldari, *et al.*, 1987. *EMBO J.* 6: 229-234), pMFa (Kurjan and Herskowitz, 1982. *Cell* 30: 933-943), pJRY88 (Schultz *et al.*, 1987. *Gene* 54: 113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (InVitrogen Corp, San Diego, Calif.).

Alternatively, NOVX can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, SF9 cells) include the pAc series (Smith, *et al.*, 1983. *Mol. Cell. Biol.* 3: 2156-2165) and the pVL series (Lucklow and Summers, 1989. *Virology* 170: 31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, 1987. *Nature* 329: 840) and pMT2PC (Kaufman, *et al.*, 1987. *EMBO J.* 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, adenovirus 2, cytomegalovirus, and simian virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see, *e.g.*, Chapters 16 and 17 of Sambrook, *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert, *et al.*, 1987. *Genes Dev.* 1: 268-277), lymphoid-specific promoters (Calame and Eaton, 1988. *Adv. Immunol.* 43: 235-275), in particular promoters of T cell receptors (Winoto and Baltimore, 1989. *EMBO J.* 8: 729-733) and immunoglobulins (Banerji, *et al.*, 1983. *Cell* 33: 729-740; Queen and Baltimore, 1983. *Cell* 33: 741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle, 1989. *Proc. Natl. Acad. Sci. USA* 86: 5473-5477), pancreas-specific promoters (Edlund, *et al.*, 1985. *Science* 230: 912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, *e.g.*, the murine hox promoters (Kessel and Gruss, 1990. *Science* 249: 374-379) and the □-fetoprotein promoter (Campes and Tilghman, 1989. *Genes Dev.* 3: 537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively-linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to NOVX mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA

molecule in a variety of cell types, for instance viral promoters and enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes *see, e.g.,* Weintraub, *et al.*, "Antisense RNA as a molecular tool for genetic analysis," *Reviews-Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, NOVX protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a

selectable marker can be introduced into a host cell on the same vector as that encoding NOVX or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) NOVX protein. Accordingly, the invention further provides methods for producing NOVX protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding NOVX protein has been introduced) in a suitable medium such that NOVX protein is produced. In another embodiment, the method further comprises isolating NOVX protein from the medium or the host cell.

Transgenic NOVX Animals

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which NOVX protein-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous NOVX sequences have been introduced into their genome or homologous recombinant animals in which endogenous NOVX sequences have been altered. Such animals are useful for studying the function and activity of NOVX protein and for identifying and evaluating modulators of NOVX protein activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous NOVX gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing NOVX-encoding nucleic acid into the male pronuclei of a fertilized oocyte (*e.g.*, by microinjection, retroviral infection) and allowing the oocyte to develop in a pseudopregnant female foster animal. The

human NOVX cDNA sequences SEQ ID NO: 1, 3, 4, 6, 8, or 9 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a non-human homologue of the human NOVX gene, such as a mouse NOVX gene, can be isolated based on hybridization to the human NOVX cDNA (described further *supra*) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably-linked to the NOVX transgene to direct expression of NOVX protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan, 1986. In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the NOVX transgene in its genome and expression of NOVX mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene-encoding NOVX protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of an NOVX gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the NOVX gene. The NOVX gene can be a human gene (*e.g.*, the cDNA of SEQ ID NO: 1, 3, 4, 6, 8, or 9), but more preferably, is a non-human homologue of a human NOVX gene. For example, a mouse homologue of human NOVX gene of SEQ ID NO: 1, 3, 4, 6, 8, or 9 can be used to construct a homologous recombination vector suitable for altering an endogenous NOVX gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous NOVX gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous NOVX gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous NOVX protein). In the homologous recombination vector, the altered portion of the NOVX gene is flanked at its 5'- and 3'-termini by additional nucleic acid of the NOVX gene to allow for homologous recombination to occur between the exogenous NOVX gene carried by the vector and an endogenous NOVX gene in an embryonic stem cell. The additional flanking

NOVX nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5'- and 3'-termini) are included in the vector. *See, e.g.,* Thomas, *et al.*, 1987. *Cell* 51: 503 for a description of homologous recombination vectors. The vector is ten introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced NOVX gene has homologously-recombined with the endogenous NOVX gene are selected. *See, e.g.*, Li, *et al.*, 1992. *Cell* 69: 915.

The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras. *See, e.g.*, Bradley, 1987. In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously-recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously-recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley, 1991. *Curr. Opin. Biotechnol.* 2: 823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, *See, e.g.*, Lakso, *et al.*, 1992. *Proc. Natl. Acad. Sci. USA* 89: 6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae. See,* O'Gorman, *et al.*, 1991. *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, *et al.*, 1997. *Nature* 385: 810-813. In brief, a cell (*e.g.*, a somatic cell) from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_0 phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to

morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell (*e.g.*, the somatic cell) is isolated.

Pharmaceutical Compositions

The NOVX nucleic acid molecules, NOVX proteins, and anti-NOVX antibodies (also referred to herein as "active compounds") of the invention, and derivatives, fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, finger's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (*i.e.*, topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid (EDTA); buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, an NOVX protein or anti-NOVX antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and

the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for

the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (*see, e.g.,* U.S. Patent No. 5,328,470) or by stereotactic injection (*see, e.g.*, Chen, *et al.*, 1994. *Proc. Natl. Acad. Sci. USA* 91: 3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

Screening and Detection Methods

The isolated nucleic acid molecules of the invention can be used to express NOVX protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect NOVX mRNA (*e.g.*, in a biological sample) or a genetic lesion in an NOVX gene, and to modulate NOVX activity, as described further, below. In addition, the NOVX proteins can be used to screen drugs or compounds that modulate the NOVX protein activity or expression as well as to treat disorders characterized by insufficient or excessive production of NOVX protein or production of NOVX protein forms that have decreased or aberrant activity compared to NOVX wild-type protein (*e.g.*; diabetes (regulates insulin release); obesity (binds and transport lipids); metabolic disturbances associated with obesity, the metabolic syndrome X as well as anorexia and wasting disorders associated with chronic diseases and various cancers, and infectious disease(possesses anti-microbial activity) and the various dyslipidemias. In addition, the anti-NOVX antibodies of the invention can be used to detect and isolate NOVX proteins and modulate NOVX activity. In yet a further aspect, the invention can be used in methods to influence appetite, absorption of nutrients and the disposition of metabolic substrates in both a positive and negative fashion.

The invention further pertains to novel agents identified by the screening assays described herein and uses thereof for treatments as described, *supra.*

Screening Assays

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) that bind to NOVX proteins or have a stimulatory or inhibitory effect on, *e.g.*, NOVX protein expression or NOVX protein activity. The invention also includes compounds identified in the screening assays described herein.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of the membrane-bound form of an NOVX protein or polypeptide or biologically-active portion thereof. The test compounds of the invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds. *See, e.g.,* Lam, 1997. *Anticancer Drug Design* 12: 145.

A "small molecule" as used herein, is meant to refer to a composition that has a molecular weight of less than about 5 kD and most preferably less than about 4 kD. Small molecules can be, *e.g.*, nucleic acids, peptides, polypeptides, peptidomimetics, carbohydrates, lipids or other organic or inorganic molecules. Libraries of chemical and biological mixtures, such as fungal, bacterial, or algal extracts, are known in the art and can be screened with any of the assays of the invention.

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt, *et al.*, 1993. *Proc. Natl. Acad. Sci. U.S.A.* 90: 6909; Erb, *et al.*, 1994. *Proc. Natl. Acad. Sci. U.S.A.* 91: 11422; Zuckermann, *et al.*, 1994. *J. Med. Chem.* 37: 2678; Cho, *et al.*, 1993. *Science* 261: 1303; Carrell, *et al.*, 1994. *Angew. Chem. Int. Ed. Engl.* 33: 2059; Carell, *et al.*, 1994. *Angew. Chem. Int. Ed. Engl.* 33: 2061; and Gallop, *et al.*, 1994. *J. Med. Chem.* 37: 1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten, 1992. *Biotechniques* 13: 412-421), or on beads (Lam, 1991. *Nature* 354: 82-84), on chips (Fodor, 1993. *Nature* 364: 555-556), bacteria (Ladner, U.S. Patent No. 5,223,409), spores (Ladner, U.S. Patent 5,233,409), plasmids (Cull, *et al.*, 1992. *Proc. Natl. Acad. Sci. USA* 89: 1865-1869) or on phage (Scott and Smith, 1990. *Science* 249: 386-390; Devlin, 1990. *Science* 249: 404-406;

Cwirla, *et al.*, 1990. *Proc. Natl. Acad. Sci. U.S.A.* 87: 6378-6382; Felici, 1991. *J. Mol. Biol.* 222: 301-310; Ladner, U.S. Patent No. 5,233,409.).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of NOVX protein, or a biologically-active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to an NOVX protein determined. The cell, for example, can of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the NOVX protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the NOVX protein or biologically-active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically-labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of NOVX protein, or a biologically-active portion thereof, on the cell surface with a known compound which binds NOVX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with an NOVX protein, wherein determining the ability of the test compound to interact with an NOVX protein comprises determining the ability of the test compound to preferentially bind to NOVX protein or a biologically-active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of NOVX protein, or a biologically-active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the NOVX protein or biologically-active portion thereof. Determining the ability of the test compound to modulate the activity of NOVX or a biologically-active portion thereof can be accomplished, for example, by determining the ability of the NOVX protein to bind to or interact with an NOVX target molecule. As used herein, a "target molecule" is a molecule with which an NOVX protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses an NOVX interacting protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. An NOVX target molecule can be a non-NOVX molecule or an NOVX

protein or polypeptide of the invention. In one embodiment, an NOVX target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (*e.g.* a signal generated by binding of a compound to a membrane-bound NOVX molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with NOVX.

Determining the ability of the NOVX protein to bind to or interact with an NOVX target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the NOVX protein to bind to or interact with an NOVX target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.* intracellular Ca^{2+}, diacylglycerol, IP_3, etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising an NOVX-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the invention is a cell-free assay comprising contacting an NOVX protein or biologically-active portion thereof with a test compound and determining the ability of the test compound to bind to the NOVX protein or biologically-active portion thereof. Binding of the test compound to the NOVX protein can be determined either directly or indirectly as described above. In one such embodiment, the assay comprises contacting the NOVX protein or biologically-active portion thereof with a known compound which binds NOVX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with an NOVX protein, wherein determining the ability of the test compound to interact with an NOVX protein comprises determining the ability of the test compound to preferentially bind to NOVX or biologically-active portion thereof as compared to the known compound.

In still another embodiment, an assay is a cell-free assay comprising contacting NOVX protein or biologically-active portion thereof with a test compound and determining the ability of the test compound to modulate (*e.g.* stimulate or inhibit) the activity of the NOVX protein or biologically-active portion thereof. Determining the ability of the test compound to modulate the activity of NOVX can be accomplished, for example, by determining the ability of the NOVX protein to bind to an NOVX target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test

compound to modulate the activity of NOVX protein can be accomplished by determining the ability of the NOVX protein further modulate an NOVX target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as described, *supra.*

In yet another embodiment, the cell-free assay comprises contacting the NOVX protein or biologically-active portion thereof with a known compound which binds NOVX protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with an NOVX protein, wherein determining the ability of the test compound to interact with an NOVX protein comprises determining the ability of the NOVX protein to preferentially bind to or modulate the activity of an NOVX target molecule.

The cell-free assays of the invention are amenable to use of both the soluble form or the membrane-bound form of NOVX protein. In the case of cell-free assays comprising the membrane-bound form of NOVX protein, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of NOVX protein is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)$_n$, N-dodecyl--N,N-dimethyl-3-ammonio-1-propane sulfonate, 3-(3-cholamidopropyl) dimethylamminiol-1-propane sulfonate (CHAPS), or 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-propane sulfonate (CHAPSO).

In more than one embodiment of the above assay methods of the invention, it may be desirable to immobilize either NOVX protein or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to NOVX protein, or interaction of NOVX protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For example, GST-NOVX fusion proteins or GST-target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or NOVX protein, and the mixture is incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to

remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described, *supra*. Alternatively, the complexes can be dissociated from the matrix, and the level of NOVX protein binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either the NOVX protein or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated NOVX protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well-known within the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with NOVX protein or target molecules, but which do not interfere with binding of the NOVX protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or NOVX protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the NOVX protein or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the NOVX protein or target molecule.

In another embodiment, modulators of NOVX protein expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of NOVX mRNA or protein in the cell is determined. The level of expression of NOVX mRNA or protein in the presence of the candidate compound is compared to the level of expression of NOVX mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of NOVX mRNA or protein expression based upon this comparison. For example, when expression of NOVX mRNA or protein is greater (*i.e.*, statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of NOVX mRNA or protein expression. Alternatively, when expression of NOVX mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of NOVX mRNA or protein expression. The level of NOVX mRNA or protein expression in the cells can be determined by methods described herein for detecting NOVX mRNA or protein.

In yet another aspect of the invention, the NOVX proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (*see, e.g.*, U.S. Patent No. 5,283,317; Zervos, *et al.*, 1993. *Cell* 72: 223-232; Madura, *et al.*, 1993. *J. Biol. Chem.* 268: 12046-12054; Bartel, *et al.*,

1993. *Biotechniques* 14: 920-924; Iwabuchi, *et al.*, 1993. *Oncogene* 8: 1693-1696; and Brent WO 94/10300), to identify other proteins that bind to or interact with NOVX ("NOVX-binding proteins" or "NOVX-bp") and modulate NOVX activity. Such NOVX-binding proteins are also likely to be involved in the propagation of signals by the NOVX proteins as, for example, upstream or downstream elements of the NOVX pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for NOVX is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming an NOVX-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) that is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene that encodes the protein which interacts with NOVX.

The invention further pertains to novel agents identified by the aforementioned screening assays and uses thereof for treatments as described herein.

Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. By way of example, and not of limitation, these sequences can be used to: (*i*) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (*ii*) identify an individual from a minute biological sample (tissue typing); and (*iii*) aid in forensic identification of a biological sample. Some of these applications are described in the subsections, below.

Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the NOVX sequences, SEQ ID NO: 1, 3, 4, 6, 8, or 9, or fragments or derivatives thereof, can be used to map the location of

the NOVX genes, respectively, on a chromosome. The mapping of the NOVX sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, NOVX genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the NOVX sequences. Computer analysis of the NOVX, sequences can be used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the NOVX sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (*e.g.*, human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but in which human cells can, the one human chromosome that contains the gene encoding the needed enzyme will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human chromosomes. *See, e.g.*, D'Eustachio, *et al.*, 1983. *Science* 220: 919-924. Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the NOVX sequences to design oligonucleotide primers, sub-localization can be achieved with panels of fragments from specific chromosomes.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical like colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more

preferably 2,000 bases, will suffice to get good results at a reasonable amount of time. For a review of this technique, *see,* Verma, *et al.*, HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, *e.g.,* in McKusick, MENDELIAN INHERITANCE IN MAN, available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, *e.g.,* Egeland, *et al.*, 1987. *Nature*, 325: 783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the NOVX gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

Tissue Typing

The NOVX sequences of the invention can also be used to identify individuals from minute biological samples. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. The sequences of the invention are useful as additional DNA markers for RFLP ("restriction fragment length polymorphisms," described in U.S. Patent No. 5,272,057).

Furthermore, the sequences of the invention can be used to provide an alternative technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the NOVX sequences described herein can be used to prepare two

PCR primers from the 5'- and 3'-termini of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the invention can be used to obtain such identification sequences from individuals and from tissue. The NOVX sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Much of the allelic variation is due to single nucleotide polymorphisms (SNPs), which include restriction fragment length polymorphisms (RFLPs).

Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers that each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO: 1, 3, 4, 6, 8, or 9 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

Predictive Medicine

The invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the invention relates to diagnostic assays for determining NOVX protein and nucleic acid expression as well as NOVX activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant NOVX expression or activity. The disorders include metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of

developing a disorder associated with NOVX protein, nucleic acid expression or activity. For example, mutations in an NOVX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with NOVX protein, nucleic acid expression, or biological activity.

Another aspect of the invention provides methods for determining NOVX protein, nucleic acid expression or activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (*e.g.*, drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (*e.g.*, the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of NOVX in clinical trials.

These and other agents are described in further detail in the following sections.

Diagnostic Assays

An exemplary method for detecting the presence or absence of NOVX in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting NOVX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes NOVX protein such that the presence of NOVX is detected in the biological sample. An agent for detecting NOVX mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to NOVX mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length NOVX nucleic acid, such as the nucleic acid of SEQ ID NO: 1, 3, 4, 6, 8, or 9, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to NOVX mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

An agent for detecting NOVX protein is an antibody capable of binding to NOVX protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a

primary antibody using a fluorescently-labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently-labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect NOVX mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of NOVX mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of NOVX protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations, and immunofluorescence. *In vitro* techniques for detection of NOVX genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of NOVX protein include introducing into a subject a labeled anti-NOVX antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting NOVX protein, mRNA, or genomic DNA, such that the presence of NOVX protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of NOVX protein, mRNA or genomic DNA in the control sample with the presence of NOVX protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of NOVX in a biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting NOVX protein or mRNA in a biological sample; means for determining the amount of NOVX in the sample; and means for comparing the amount of NOVX in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect NOVX protein or nucleic acid.

Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant NOVX expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a

disorder associated with NOVX protein, nucleic acid expression or activity. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the invention provides a method for identifying a disease or disorder associated with aberrant NOVX expression or activity in which a test sample is obtained from a subject and NOVX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) is detected, wherein the presence of NOVX protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant NOVX expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant NOVX expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder. Thus, the invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant NOVX expression or activity in which a test sample is obtained and NOVX protein or nucleic acid is detected (*e.g.*, wherein the presence of NOVX protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant NOVX expression or activity).

The methods of the invention can also be used to detect genetic lesions in an NOVX gene, thereby determining if a subject with the lesioned gene is at risk for a disorder characterized by aberrant cell proliferation and differentiation. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding an NOVX-protein, or the misexpression of the NOVX gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of: (*i*) a deletion of one or more nucleotides from an NOVX gene; (*ii*) an addition of one or more nucleotides to an NOVX gene; (*iii*) a substitution of one or more nucleotides of an NOVX gene, (*iv*) a chromosomal rearrangement of an NOVX gene; (*v*) an alteration in the level of a messenger RNA transcript of an NOVX gene, (*vi*) aberrant modification of an NOVX gene, such as of the methylation pattern of the genomic DNA, (*vii*) the presence of a non-wild-type splicing pattern of a messenger RNA transcript of an NOVX gene, (*viii*) a non-wild-type level of an NOVX protein, (*ix*) allelic loss of an NOVX gene, and (*x*) inappropriate post-translational modification of an NOVX protein. As described herein, there are a large number of assay techniques known

in the art which can be used for detecting lesions in an NOVX gene. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (*see, e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (*see, e.g.*, Landegran, *et al.*, 1988. *Science* 241: 1077-1080; and Nakazawa, *et al.*, 1994. *Proc. Natl. Acad. Sci. USA* 91: 360-364), the latter of which can be particularly useful for detecting point mutations in the NOVX-gene (*see,* Abravaya, *et al.*, 1995. *Nucl. Acids Res.* 23: 675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers that specifically hybridize to an NOVX gene under conditions such that hybridization and amplification of the NOVX gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (*see,* Guatelli, *et al.*, 1990. *Proc. Natl. Acad. Sci. USA* 87: 1874-1878), transcriptional amplification system (*see,* Kwoh, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA* 86: 1173-1177); Qβ Replicase (*see,* Lizardi, *et al*, 1988. *BioTechnology* 6: 1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in an NOVX gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (*see, e.g.,* U.S. Patent No. 5,493,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in NOVX can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high-density arrays containing hundreds or thousands of oligonucleotides probes. *See, e.g.,* Cronin, *et al.*, 1996. *Human Mutation* 7: 244-255; Kozal, *et al.*, 1996. *Nat. Med.* 2: 753-759. For example, genetic mutations in NOVX can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, *et al.*, *supra.* Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the NOVX gene and detect mutations by comparing the sequence of the sample NOVX with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert, 1977. *Proc. Natl. Acad. Sci. USA* 74: 560 or Sanger, 1977. *Proc. Natl. Acad. Sci. USA* 74: 5463. It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays (*see, e.g.,* Naeve, *et al.*, 1995. *Biotechniques* 19: 448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publication No. WO 94/16101; Cohen, *et al.*, 1996. *Adv. Chromatography* 36: 127-162; and Griffin, *et al.*, 1993. *Appl. Biochem. Biotechnol.* 38: 147-159).

Other methods for detecting mutations in the NOVX gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes. *See, e.g.,* Myers, *et al.*, 1985. *Science* 230: 1242. In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type NOVX sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent that cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S_1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with

piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. *See, e.g.,* Cotton, *et al.,* 1988. *Proc. Natl. Acad. Sci. USA* 85: 4397; Saleeba, *et al.,* 1992. *Methods Enzymol.* 217: 286-295. In an embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in NOVX cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches. *See, e.g.,* Hsu, *et al.,* 1994. *Carcinogenesis* 15: 1657-1662. According to an exemplary embodiment, a probe based on an NOVX sequence, *e.g.,* a wild-type NOVX sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. *See, e.g.,* U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in NOVX genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids. *See, e.g.,* Orita, *et al.,* 1989. *Proc. Natl. Acad. Sci. USA*: 86: 2766; Cotton, 1993. *Mutat. Res.* 285: 125-144; Hayashi, 1992. *Genet. Anal. Tech. Appl.* 9: 73-79. Single-stranded DNA fragments of sample and control NOVX nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility. *See, e.g.,* Keen, *et al.,* 1991. *Trends Genet.* 7: 5.

In yet another embodiment, the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE). *See, e.g.,* Myers, *et al.,* 1985. *Nature* 313: 495. When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich

DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA. *See, e.g.,* Rosenbaum and Reissner, 1987. *Biophys. Chem.* 265: 12753.

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions that permit hybridization only if a perfect match is found. *See, e.g.,* Saiki, *et al.*, 1986. *Nature* 324: 163; Saiki, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA* 86: 6230. Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization; *see, e.g.,* Gibbs, *et al.*, 1989. *Nucl. Acids Res.* 17: 2437-2448) or at the extreme 3'-terminus of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (*see, e.g.,* Prossner, 1993. *Tibtech.* 11: 238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection. *See, e.g.,* Gasparini, *et al.*, 1992. *Mol. Cell Probes* 6: 1. It is anticipated that in certain embodiments amplification may also be performed using *Taq* ligase for amplification. *See, e.g.,* Barany, 1991. *Proc. Natl. Acad. Sci. USA* 88: 189. In such cases, ligation will occur only if there is a perfect match at the 3'-terminus of the 5' sequence, making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving an NOVX gene.

Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which NOVX is expressed may be utilized in the prognostic assays described herein. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

Pharmacogenomics

Agents, or modulators that have a stimulatory or inhibitory effect on NOVX activity (*e.g.*, NOVX gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (The disorders include metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers.) In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of NOVX protein, expression of NOVX nucleic acid, or mutation content of NOVX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See *e.g.*, Eichelbaum, 1996. *Clin. Exp. Pharmacol. Physiol.*, 23: 983-985; Linder, 1997. *Clin. Chem.*, 43: 254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is hemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and

cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. At the other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of NOVX protein, expression of NOVX nucleic acid, or mutation content of NOVX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with an NOVX modulator, such as a modulator identified by one of the exemplary screening assays described herein.

Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of NOVX (*e.g.*, the ability to modulate aberrant cell proliferation and differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase NOVX gene expression, protein levels, or upregulate NOVX activity, can be monitored in clinical trails of subjects exhibiting decreased NOVX gene expression, protein levels, or downregulated NOVX activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease NOVX gene expression, protein levels, or downregulate NOVX activity, can be monitored in clinical trails of subjects exhibiting increased NOVX gene expression, protein levels, or upregulated NOVX activity. In such clinical trials, the expression or activity of NOVX and, preferably, other genes that have been implicated in, for example, a

cellular proliferation or immune disorder can be used as a "read out" or markers of the immune responsiveness of a particular cell.

By way of example, and not of limitation, genes, including NOVX, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) that modulates NOVX activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of NOVX and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of NOVX or other genes. In this manner, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, peptidomimetic, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (*i*) obtaining a pre-administration sample from a subject prior to administration of the agent; (*ii*) detecting the level of expression of an NOVX protein, mRNA, or genomic DNA in the preadministration sample; (*iii*) obtaining one or more post-administration samples from the subject; (*iv*) detecting the level of expression or activity of the NOVX protein, mRNA, or genomic DNA in the post-administration samples; (*v*) comparing the level of expression or activity of the NOVX protein, mRNA, or genomic DNA in the pre-administration sample with the NOVX protein, mRNA, or genomic DNA in the post administration sample or samples; and (*vi*) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of NOVX to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of NOVX to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

Methods of Treatment

The invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant NOVX expression or activity. The disorders include cardiomyopathy, atherosclerosis,

hypertension, congenital heart defects, aortic stenosis, atrial septal defect (ASD), atrioventricular (A-V) canal defect, ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular septal defect (VSD), valve diseases, tuberous sclerosis, scleroderma, obesity, transplantation, adrenoleukodystrophy, congenital adrenal hyperplasia, prostate cancer, neoplasm; adenocarcinoma, lymphoma, uterus cancer, fertility, hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, immunodeficiencies, graft versus host disease, AIDS, bronchial asthma, Crohn's disease; multiple sclerosis, treatment of Albright Hereditary Ostoeodystrophy, and other diseases, disorders and conditions of the like.

These methods of treatment will be discussed more fully, below.

Disease and Disorders

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to: (*i*) an aforementioned peptide, or analogs, derivatives, fragments or homologs thereof; (*ii*) antibodies to an aforementioned peptide; (*iii*) nucleic acids encoding an aforementioned peptide; (*iv*) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to an aforementioned peptide) that are utilized to "knockout" endogenous function of an aforementioned peptide by homologous recombination (*see*, *e.g.*, Capecchi, 1989. *Science* 244: 1288-1292); or (*v*) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between an aforementioned peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, an aforementioned peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and activity of the expressed peptides (or mRNAs of an aforementioned peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by

sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, and the like).

Prophylactic Methods

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant NOVX expression or activity, by administering to the subject an agent that modulates NOVX expression or at least one NOVX activity. Subjects at risk for a disease that is caused or contributed to by aberrant NOVX expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the NOVX aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending upon the type of NOVX aberrancy, for example, an NOVX agonist or NOVX antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein. The prophylactic methods of the invention are further discussed in the following subsections.

Therapeutic Methods

Another aspect of the invention pertains to methods of modulating NOVX expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of NOVX protein activity associated with the cell. An agent that modulates NOVX protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of an NOVX protein, a peptide, an NOVX peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more NOVX protein activity. Examples of such stimulatory agents include active NOVX protein and a nucleic acid molecule encoding NOVX that has been introduced into the cell. In another embodiment, the agent inhibits one or more NOVX protein activity. Examples of such inhibitory agents include antisense NOVX nucleic acid molecules and anti-NOVX antibodies. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of an NOVX protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, up-regulates or down-regulates) NOVX expression or activity. In another

embodiment, the method involves administering an NOVX protein or nucleic acid molecule as therapy to compensate for reduced or aberrant NOVX expression or activity.

Stimulation of NOVX activity is desirable *in situ*ations in which NOVX is abnormally downregulated and in which increased NOVX activity is likely to have a beneficial effect. One example of such a situation is where a subject has a disorder characterized by aberrant cell proliferation and differentiation (*e.g.*, cancer or immune associated disorders). Another example of such a situation is where the subject has a gestational disease (*e.g.*, preclampsia).

Determination of the Biological Effect of the Therapeutic

In various embodiments of the invention, suitable *in vitro* or *in vivo* assays are performed to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue.

In various specific embodiments, *in vitro* assays may be performed with representative cells of the type(s) involved in the patient's disorder, to determine if a given Therapeutic exerts the desired effect upon the cell type(s). Compounds for use in therapy may be tested in suitable animal model systems including, but not limited to rats, mice, chicken, cows, monkeys, rabbits, and the like, prior to testing in human subjects. Similarly, for *in vivo* testing, any of the animal model system known in the art may be used prior to administration to human subjects.

Prophylactic and Therapeutic Uses of the Compositions of the Invention

The NOVX nucleic acids and proteins of the invention are useful in potential prophylactic and therapeutic applications implicated in a variety of disorders including, but not limited to: metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers.

As an example, a cDNA encoding the NOVX protein of the invention may be useful in gene therapy, and the protein may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the invention will have efficacy for treatment of patients suffering from: metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, hematopoietic disorders, and the various dyslipidemias.

Both the novel nucleic acid encoding the NOVX protein, and the NOVX protein of the invention, or fragments thereof, may also be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. A further use could be as an anti-bacterial molecule (*i.e.,* some peptides have been found to possess anti-bacterial properties). These materials are further useful in the generation of antibodies, which immunospecifically-bind to the novel substances of the invention for use in therapeutic or diagnostic methods.

The invention will be further described in the following examples, which do not limit the scope of the invention described in the claims.

Examples

Example 1. Identification of NOV1

SeqCallingTM Technology: cDNA was derived from various human samples representing multiple tissue types, normal and diseased states, physiological states, and developmental states from different donors. Samples were obtained as whole tissue, cell lines, primary cells or tissue cultured primary cells and cell lines. Cells and cell lines may have been treated with biological or chemical agents that regulate gene expression for example, growth factors, chemokines, steroids. The cDNA thus derived was then sequenced using CuraGen's proprietary SeqCalling technology. Sequence traces were evaluated manually and edited for corrections if appropriate. cDNA sequences from all samples were assembled with themselves and with public ESTs using bioinformatics programs to generate CuraGen's human SeqCalling database of SeqCalling assemblies. Each assembly contains one or more overlapping cDNA sequences derived from one or more human samples. Fragments and ESTs were included as components for an assembly when the extent of identity with another component of the assembly was at least 95% over 50 bp. Each assembly can represent a gene and its variants such as splice forms and single nucleotide polymorphisms (SNPs) and their combinations.

Variant sequences are included in this application. A variant sequence can include a single nucleotide polymorphism (SNP). A SNP can, in some instances, be referred to as a "cSNP" to denote that the nucleotide sequence containing the SNP originates as a cDNA. A SNP can arise in several ways. For example, a SNP may be due to a substitution of one nucleotide for another at the polymorphic site. Such a substitution can be either a transition or a transversion. A SNP can also arise from a deletion of a nucleotide or an insertion of a nucleotide, relative to a reference allele. In this case, the polymorphic site is a site at which one

allele bears a gap with respect to a particular nucleotide in another allele. SNPs occurring within genes may result in an alteration of the amino acid encoded by the gene at the position of the SNP. Intragenic SNPs may also be silent, however, in the case that a codon including a SNP encodes the same amino acid as a result of the redundancy of the genetic code. SNPs occurring outside the region of a gene, or in an intron within a gene, do not result in changes in any amino acid sequence of a protein but may result in altered regulation of the expression pattern for example, alteration in temporal expression, physiological response regulation, cell type expression regulation, intensity of expression, stability of transcribed message.

Example 2. Identification of NOV2

The novel NOV2 target sequence identified in the present invention were subjected to the exon linking process to confirm the sequence. Table 3A shows the sequences of the PCR primers used for obtaining different clones. Such primers were designed based on in silico predictions for the full length cDNA, part (one or more exons) of the DNA or protein sequence of the target sequence, or by translated homology of the predicted exons to closely related human sequences from other species. These primers were then employed in PCR amplification based on the following pool of human cDNAs: adrenal gland, bone marrow, brain - amygdala, brain - cerebellum, brain - hippocampus, brain - substantia nigra, brain - thalamus, brain - whole, fetal brain, fetal kidney, fetal liver, fetal lung, heart, kidney, lymphoma - Raji, mammary gland, pancreas, pituitary gland, placenta, prostate, salivary gland, skeletal muscle, small intestine, spinal cord, spleen, stomach, testis, thyroid, trachea, and uterus. The PCR product derived from exon linking was cloned into the pCR2.1 vector from Invitrogen. The resulting bacterial clone has an insert covering the entire open reading frame cloned into the pCR2.1 vector. Table 3B shows a list of these bacterial clones. The resulting sequences from all clones were assembled with themselves, with other fragments in CuraGen Corporation's database and with public ESTs. Fragments and ESTs were included as components for an assembly when the extent of their identity with another component of the assembly was at least 95% over 50 bp. In addition, sequence traces were evaluated manually and edited for corrections if appropriate. These procedures provide the sequence reported herein.

Table 3A. PCR Primers for Exon Linking

NOVX Clone	Primer 1 (5′ - 3′)	SEQ ID NO	Primer 2 (5′ - 3′)	SEQ ID NO
NOV2	CCTTTTGGGGCATGTTGATCCG	16	CAGTTCCAGTAGTCTTCTTGAGGAAAACACCA	17

Physical clone: Exons were predicted by homology and the intron/exon boundaries were determined using standard genetic rules. Exons were further selected and refined by means of similarity determination using multiple BLAST (for example, tBlastN, BlastX, and BlastN) searches, and, in some instances, GeneScan and Grail. Expressed sequences from both public and proprietary databases were also added when available to further define and complete the gene sequence. The DNA sequence was then manually corrected for apparent inconsistencies thereby obtaining the sequences encoding the full-length protein.

Table 3B. Physical Clones for PCR products

NOVX Clone	Bacterial Clone
NOV2	20936375_0_1_da2.69832.L4

Example 3. SNP analyses of NOVX clones

Method of novel SNP Identification: SNPs are identified by analyzing sequence assemblies using CuraGen's proprietary SNPTool algorithm. SNPTool identifies variation in assemblies with the following criteria: SNPs are not analyzed within 10 base pairs on both ends of an alignment; Window size (number of bases in a view) is 10; The allowed number of mismatches in a window is 2; Minimum SNP base quality (PHRED score) is 23; Minimum number of changes to score an SNP is 2/assembly position. SNPTool analyzes the assembly and displays SNP positions, associated individual variant sequences in the assembly, the depth of the assembly at that given position, the putative assembly allele frequency, and the SNP sequence variation. Sequence traces are then selected and brought into view for manual validation. The consensus assembly sequence is imported into CuraTools along with variant sequence changes to identify potential amino acid changes resulting from the SNP sequence variation. Comprehensive SNP data analysis is then exported into the SNPCalling database.

Method of novel SNP Confirmation: SNPs are confirmed employing a validated method know as Pyrosequencing (Pyrosequencing, Westborough, MA). Detailed protocols for Pyrosequencing can be found in: Alderborn et al. Determination of Single Nucleotide Polymorphisms by Real-time Pyrophosphate DNA Sequencing. (2000). Genome Research. 10, Issue 8, August. 1249-1265. In brief, Pyrosequencing is a real time primer extension process of genotyping. This protocol takes double-stranded, biotinylated PCR products from genomic DNA samples and binds them to streptavidin beads. These beads are then denatured producing single stranded bound DNA. SNPs are characterized utilizing a technique based on an indirect bioluminometric assay of pyrophosphate (PPi) that is released from each dNTP upon DNA chain elongation. Following Klenow polymerase-mediated base incorporation, PPi is released

and used as a substrate, together with adenosine 5'-phosphosulfate (APS), for ATP sulfurylase, which results in the formation of ATP. Subsequently, the ATP accomplishes the conversion of luciferin to its oxi-derivative by the action of luciferase. The ensuing light output becomes proportional to the number of added bases, up to about four bases. To allow processivity of the method dNTP excess is degraded by apyrase, which is also present in the starting reaction mixture, so that only dNTPs are added to the template during the sequencing. The process has been fully automated and adapted to a 96-well format, which allows rapid screening of large SNP panels.

Example 4. Quantitative expression analysis of clones in various tissues and cells

The quantitative expression of various clones was assessed using microtiter plates containing RNA samples from a variety of normal and pathology-derived cells, cell lines and tissues using real time quantitative PCR (RTQ PCR). RTQ PCR was performed on an Applied Biosystems ABI PRISM® 7700 or an ABI PRISM® 7900 HT Sequence Detection System. Various collections of samples are assembled on the plates, and referred to as Panel 1 (containing normal tissues and cancer cell lines), Panel 2 (containing samples derived from tissues from normal and cancer sources), Panel 3 (containing cancer cell lines), Panel 4 (containing cells and cell lines from normal tissues and cells related to inflammatory conditions), Panel 5D/5I (containing human tissues and cell lines with an emphasis on metabolic diseases), AI_comprehensive_panel (containing normal tissue and samples from autoinflammatory diseases), Panel CNSD.01 (containing samples from normal and diseased brains) and CNS_neurodegeneration_panel (containing samples from normal and Alzheimer's diseased brains).

RNA integrity from all samples is controlled for quality by visual assessment of agarose gel electropherograms using 28S and 18S ribosomal RNA staining intensity ratio as a guide (2:1 to 2.5:1 28s:18s) and the absence of low molecular weight RNAs that would be indicative of degradation products. Samples are controlled against genomic DNA contamination by RTQ PCR reactions run in the absence of reverse transcriptase using probe and primer sets designed to amplify across the span of a single exon.

First, the RNA samples were normalized to reference nucleic acids such as constitutively expressed genes (for example, β-actin and GAPDH). Normalized RNA (5 ul) was converted to cDNA and analyzed by RTQ-PCR using One Step RT-PCR Master Mix Reagents (Applied Biosystems; Catalog No. 4309169) and gene-specific primers according to the manufacturer's instructions.

In other cases, non-normalized RNA samples were converted to single strand cDNA (sscDNA) using Superscript II (Invitrogen Corporation; Catalog No. 18064-147) and random hexamers according to the manufacturer's instructions. Reactions containing up to 10 µg of total RNA were performed in a volume of 20 µl and incubated for 60 minutes at 42°C. This reaction can be scaled up to 50 µg of total RNA in a final volume of 100 µl. sscDNA samples are then normalized to reference nucleic acids as described previously, using 1X TaqMan® Universal Master mix (Applied Biosystems; catalog No. 4324020), following the manufacturer's instructions.

Probes and primers were designed for each assay according to Applied Biosystems Primer Express Software package (version I for Apple Computer's Macintosh Power PC) or a similar algorithm using the target sequence as input. Default settings were used for reaction conditions and the following parameters were set before selecting primers: primer concentration = 250 nM, primer melting temperature (Tm) range = 58°-60°C, primer optimal Tm = 59°C, maximum primer difference = 2°C, probe does not have 5'G, probe Tm must be 10°C greater than primer Tm, amplicon size 75bp to 100bp. The probes and primers selected (see below) were synthesized by Synthegen (Houston, TX, USA). Probes were double purified by HPLC to remove uncoupled dye and evaluated by mass spectroscopy to verify coupling of reporter and quencher dyes to the 5' and 3' ends of the probe, respectively. Their final concentrations were: forward and reverse primers, 900nM each, and probe, 200nM.

PCR conditions: When working with RNA samples, normalized RNA from each tissue and each cell line was spotted in each well of either a 96 well or a 384-well PCR plate (Applied Biosystems). PCR cocktails included either a single gene specific probe and primers set, or two multiplexed probe and primers sets (a set specific for the target clone and another gene-specific set multiplexed with the target probe). PCR reactions were set up using TaqMan® One-Step RT-PCR Master Mix (Applied Biosystems, Catalog No. 4313803) following manufacturer's instructions. Reverse transcription was performed at 48°C for 30 minutes followed by amplification/PCR cycles as follows: 95°C 10 min, then 40 cycles of 95°C for 15 seconds, 60°C for 1 minute. Results were recorded as CT values (cycle at which a given sample crosses a threshold level of fluorescence) using a log scale, with the difference in RNA concentration between a given sample and the sample with the lowest CT value being represented as 2 to the power of delta CT. The percent relative expression is then obtained by taking the reciprocal of this RNA difference and multiplying by 100.

When working with sscDNA samples, normalized sscDNA was used as described previously for RNA samples. PCR reactions containing one or two sets of probe and primers

were set up as described previously, using 1X TaqMan® Universal Master mix (Applied Biosystems; catalog No. 4324020), following the manufacturer's instructions. PCR amplification was performed as follows: 95°C 10 min, then 40 cycles of 95°C for 15 seconds, 60°C for 1 minute. Results were analyzed and processed as described previously.

Panels 1, 1.1, 1.2, and 1.3D

The plates for Panels 1, 1.1, 1.2 and 1.3D include 2 control wells (genomic DNA control and chemistry control) and 94 wells containing cDNA from various samples. The samples in these panels are broken into 2 classes: samples derived from cultured cell lines and samples derived from primary normal tissues. The cell lines are derived from cancers of the following types: lung cancer, breast cancer, melanoma, colon cancer, prostate cancer, CNS cancer, squamous cell carcinoma, ovarian cancer, liver cancer, renal cancer, gastric cancer and pancreatic cancer. Cell lines used in these panels are widely available through the American Type Culture Collection (ATCC), a repository for cultured cell lines, and were cultured using the conditions recommended by the ATCC. The normal tissues found on these panels are comprised of samples derived from all major organ systems from single adult individuals or fetuses. These samples are derived from the following organs: adult skeletal muscle, fetal skeletal muscle, adult heart, fetal heart, adult kidney, fetal kidney, adult liver, fetal liver, adult lung, fetal lung, various regions of the brain, the spleen, bone marrow, lymph node, pancreas, salivary gland, pituitary gland, adrenal gland, spinal cord, thymus, stomach, small intestine, colon, bladder, trachea, breast, ovary, uterus, placenta, prostate, testis and adipose.

In the results for Panels 1, 1.1, 1.2 and 1.3D, the following abbreviations are used:

ca. = carcinoma,
* = established from metastasis,
met = metastasis,
s cell var = small cell variant,
non-s = non-sm = non-small,
squam = squamous,
pl. eff = pl effusion = pleural effusion,
glio = glioma,
astro = astrocytoma, and
neuro = neuroblastoma.

General_screening_panel_v1.4

The plates for Panel 1.4 include 2 control wells (genomic DNA control and chemistry control) and 94 wells containing cDNA from various samples. The samples in Panel 1.4 are broken into 2 classes: samples derived from cultured cell lines and samples derived from

primary normal tissues. The cell lines are derived from cancers of the following types: lung cancer, breast cancer, melanoma, colon cancer, prostate cancer, CNS cancer, squamous cell carcinoma, ovarian cancer, liver cancer, renal cancer, gastric cancer and pancreatic cancer. Cell lines used in Panel 1.4 are widely available through the American Type Culture Collection (ATCC), a repository for cultured cell lines, and were cultured using the conditions recommended by the ATCC. The normal tissues found on Panel 1.4 are comprised of pools of samples derived from all major organ systems from 2 to 5 different adult individuals or fetuses.

These samples are derived from the following organs: adult skeletal muscle, fetal skeletal muscle, adult heart, fetal heart, adult kidney, fetal kidney, adult liver, fetal liver, adult lung, fetal lung, various regions of the brain, the spleen, bone marrow, lymph node, pancreas, salivary gland, pituitary gland, adrenal gland, spinal cord, thymus, stomach, small intestine, colon, bladder, trachea, breast, ovary, uterus, placenta, prostate, testis and adipose. Abbreviations are as described for Panels 1, 1.1, 1.2, and 1.3D.

Panels 2D and 2.2

The plates for Panels 2D and 2.2 generally include 2 control wells and 94 test samples composed of RNA or cDNA isolated from human tissue procured by surgeons working in close cooperation with the National Cancer Institute's Cooperative Human Tissue Network (CHTN) or the National Disease Research Initiative (NDRI). The tissues are derived from human malignancies and in cases where indicated many malignant tissues have "matched margins" obtained from noncancerous tissue just adjacent to the tumor. These are termed normal adjacent tissues and are denoted "NAT" in the results below. The tumor tissue and the "matched margins" are evaluated by two independent pathologists (the surgical pathologists and again by a pathologist at NDRI or CHTN). This analysis provides a gross histopathological assessment of tumor differentiation grade. Moreover, most samples include the original surgical pathology report that provides information regarding the clinical stage of the patient. These matched margins are taken from the tissue surrounding (i.e. immediately proximal) to the zone of surgery (designated "NAT", for normal adjacent tissue, in Table RR). In addition, RNA and cDNA samples were obtained from various human tissues derived from autopsies performed on elderly people or sudden death victims (accidents, etc.). These tissues were ascertained to be free of disease and were purchased from various commercial sources such as Clontech (Palo Alto, CA), Research Genetics, and Invitrogen.

Panel 3D

The plates of Panel 3D are comprised of 94 cDNA samples and two control samples. Specifically, 92 of these samples are derived from cultured human cancer cell lines, 2 samples of human primary cerebellar tissue and 2 controls. The human cell lines are generally obtained from ATCC (American Type Culture Collection), NCI or the German tumor cell bank and fall into the following tissue groups: Squamous cell carcinoma of the tongue, breast cancer, prostate cancer, melanoma, epidermoid carcinoma, sarcomas, bladder carcinomas, pancreatic cancers, kidney cancers, leukemias/lymphomas, ovarian/uterine/cervical, gastric, colon, lung and CNS cancer cell lines. In addition, there are two independent samples of cerebellum. These cells are all cultured under standard recommended conditions and RNA extracted using the standard procedures. The cell lines in panel 3D and 1.3D are of the most common cell lines used in the scientific literature.

Panels 4D, 4R, and 4.1D

Panel 4 includes samples on a 96 well plate (2 control wells, 94 test samples) composed of RNA (Panel 4R) or cDNA (Panels 4D/4.1D) isolated from various human cell lines or tissues related to inflammatory conditions. Total RNA from control normal tissues such as colon and lung (Stratagene, La Jolla, CA) and thymus and kidney (Clontech) was employed. Total RNA from liver tissue from cirrhosis patients and kidney from lupus patients was obtained from BioChain (Biochain Institute, Inc., Hayward, CA). Intestinal tissue for RNA preparation from patients diagnosed as having Crohn's disease and ulcerative colitis was obtained from the National Disease Research Interchange (NDRI) (Philadelphia, PA).

Astrocytes, lung fibroblasts, dermal fibroblasts, coronary artery smooth muscle cells, small airway epithelium, bronchial epithelium, microvascular dermal endothelial cells, microvascular lung endothelial cells, human pulmonary aortic endothelial cells, human umbilical vein endothelial cells were all purchased from Clonetics (Walkersville, MD) and grown in the media supplied for these cell types by Clonetics. These primary cell types were activated with various cytokines or combinations of cytokines for 6 and 12-14 hours, as indicated. The following cytokines were used; IL-1 beta at approximately 1-5ng/ml, TNF alpha at approximately 5-10ng/ml, IFN gamma at approximately 20-50ng/ml, IL-4 at approximately 5-10ng/ml, IL-9 at approximately 5-10ng/ml, IL-13 at approximately 5-10ng/ml. Endothelial cells were sometimes starved for various times by culture in the basal media from Clonetics with 0.1% serum.

Mononuclear cells were prepared from blood of employees at CuraGen Corporation, using Ficoll. LAK cells were prepared from these cells by culture in DMEM 5% FCS

(Hyclone), 100μM non essential amino acids (Gibco/Life Technologies, Rockville, MD), 1mM sodium pyruvate (Gibco), mercaptoethanol $5.5x10^{-5}$M (Gibco), and 10mM Hepes (Gibco) and Interleukin 2 for 4-6 days. Cells were then either activated with 10-20ng/ml PMA and 1-2μg/ml ionomycin, IL-12 at 5-10ng/ml, IFN gamma at 20-50ng/ml and IL-18 at 5-10ng/ml for 6 hours. In some cases, mononuclear cells were cultured for 4-5 days in DMEM 5% FCS (Hyclone), 100μM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol $5.5x10^{-5}$M (Gibco), and 10mM Hepes (Gibco) with PHA (phytohemagglutinin) or PWM (pokeweed mitogen) at approximately 5μg/ml. Samples were taken at 24, 48 and 72 hours for RNA preparation. MLR (mixed lymphocyte reaction) samples were obtained by taking blood from two donors, isolating the mononuclear cells using Ficoll and mixing the isolated mononuclear cells 1:1 at a final concentration of approximately $2x10^{6}$cells/ml in DMEM 5% FCS (Hyclone), 100μM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol ($5.5x10^{-5}$M) (Gibco), and 10mM Hepes (Gibco). The MLR was cultured and samples taken at various time points ranging from 1- 7 days for RNA preparation.

Monocytes were isolated from mononuclear cells using CD14 Miltenyi Beads, +ve VS selection columns and a Vario Magnet according to the manufacturer's instructions. Monocytes were differentiated into dendritic cells by culture in DMEM 5% fetal calf serum (FCS) (Hyclone, Logan, UT), 100μM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol $5.5x10^{-5}$M (Gibco), and 10mM Hepes (Gibco), 50ng/ml GMCSF and 5ng/ml IL-4 for 5-7 days. Macrophages were prepared by culture of monocytes for 5-7 days in DMEM 5% FCS (Hyclone), 100μM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol $5.5x10^{-5}$M (Gibco), 10mM Hepes (Gibco) and 10% AB Human Serum or MCSF at approximately 50ng/ml. Monocytes, macrophages and dendritic cells were stimulated for 6 and 12-14 hours with lipopolysaccharide (LPS) at 100ng/ml. Dendritic cells were also stimulated with anti-CD40 monoclonal antibody (Pharmingen) at 10μg/ml for 6 and 12-14 hours.

CD4 lymphocytes, CD8 lymphocytes and NK cells were also isolated from mononuclear cells using CD4, CD8 and CD56 Miltenyi beads, positive VS selection columns and a Vario Magnet according to the manufacturer's instructions. CD45RA and CD45RO CD4 lymphocytes were isolated by depleting mononuclear cells of CD8, CD56, CD14 and CD19 cells using CD8, CD56, CD14 and CD19 Miltenyi beads and positive selection. CD45RO beads were then used to isolate the CD45RO CD4 lymphocytes with the remaining cells being CD45RA CD4 lymphocytes. CD45RA CD4, CD45RO CD4 and CD8 lymphocytes were placed in DMEM 5% FCS (Hyclone), 100μM non essential amino acids (Gibco), 1mM sodium

pyruvate (Gibco), mercaptoethanol $5.5x10^{-5}$M (Gibco), and 10mM Hepes (Gibco) and plated at 10^6cells/ml onto Falcon 6 well tissue culture plates that had been coated overnight with 0.5μg/ml anti-CD28 (Pharmingen) and 3ug/ml anti-CD3 (OKT3, ATCC) in PBS. After 6 and 24 hours, the cells were harvested for RNA preparation. To prepare chronically activated CD8 lymphocytes, we activated the isolated CD8 lymphocytes for 4 days on anti-CD28 and anti-CD3 coated plates and then harvested the cells and expanded them in DMEM 5% FCS (Hyclone), 100μM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol $5.5x10^{-5}$M (Gibco), and 10mM Hepes (Gibco) and IL-2. The expanded CD8 cells were then activated again with plate bound anti-CD3 and anti-CD28 for 4 days and expanded as before. RNA was isolated 6 and 24 hours after the second activation and after 4 days of the second expansion culture. The isolated NK cells were cultured in DMEM 5% FCS (Hyclone), 100μM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol $5.5x10^{-5}$M (Gibco), and 10mM Hepes (Gibco) and IL-2 for 4-6 days before RNA was prepared.

To obtain B cells, tonsils were procured from NDRI. The tonsil was cut up with sterile dissecting scissors and then passed through a sieve. Tonsil cells were then spun down and resupended at 10^6cells/ml in DMEM 5% FCS (Hyclone), 100μM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol $5.5x10^{-5}$M (Gibco), and 10mM Hepes (Gibco). To activate the cells, we used PWM at 5μg/ml or anti-CD40 (Pharmingen) at approximately 10μg/ml and IL-4 at 5-10ng/ml. Cells were harvested for RNA preparation at 24,48 and 72 hours.

To prepare the primary and secondary Th1/Th2 and Tr1 cells, six-well Falcon plates were coated overnight with 10μg/ml anti-CD28 (Pharmingen) and 2μg/ml OKT3 (ATCC), and then washed twice with PBS. Umbilical cord blood CD4 lymphocytes (Poietic Systems, German Town, MD) were cultured at 10^5-10^6cells/ml in DMEM 5% FCS (Hyclone), 100μM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol $5.5x10^{-5}$M (Gibco), 10mM Hepes (Gibco) and IL-2 (4ng/ml). IL-12 (5ng/ml) and anti-IL4 (1μg/ml) were used to direct to Th1, while IL-4 (5ng/ml) and anti-IFN gamma (1μg/ml) were used to direct to Th2 and IL-10 at 5ng/ml was used to direct to Tr1. After 4-5 days, the activated Th1, Th2 and Tr1 lymphocytes were washed once in DMEM and expanded for 4-7 days in DMEM 5% FCS (Hyclone), 100μM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol $5.5x10^{-5}$M (Gibco), 10mM Hepes (Gibco) and IL-2 (1ng/ml). Following this, the activated Th1, Th2 and Tr1 lymphocytes were re-stimulated for 5 days with anti-CD28/OKT3 and cytokines as described above, but with the addition of anti-CD95L (1μg/ml)

to prevent apoptosis. After 4-5 days, the Th1, Th2 and Tr1 lymphocytes were washed and then expanded again with IL-2 for 4-7 days. Activated Th1 and Th2 lymphocytes were maintained in this way for a maximum of three cycles. RNA was prepared from primary and secondary Th1, Th2 and Tr1 after 6 and 24 hours following the second and third activations with plate bound anti-CD3 and anti-CD28 mAbs and 4 days into the second and third expansion cultures in Interleukin 2.

The following leukocyte cells lines were obtained from the ATCC: Ramos, EOL-1, KU-812. EOL cells were further differentiated by culture in 0.1mM dbcAMP at $5x10^5$cells/ml for 8 days, changing the media every 3 days and adjusting the cell concentration to $5x10^5$cells/ml. For the culture of these cells, we used DMEM or RPMI (as recommended by the ATCC), with the addition of 5% FCS (Hyclone), 100µM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol $5.5x10^{-5}$M (Gibco), 10mM Hepes (Gibco). RNA was either prepared from resting cells or cells activated with PMA at 10ng/ml and ionomycin at 1µg/ml for 6 and 14 hours. Keratinocyte line CCD106 and an airway epithelial tumor line NCI-H292 were also obtained from the ATCC. Both were cultured in DMEM 5% FCS (Hyclone), 100µM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol $5.5x10^{-5}$M (Gibco), and 10mM Hepes (Gibco). CCD1106 cells were activated for 6 and 14 hours with approximately 5 ng/ml TNF alpha and 1ng/ml IL-1 beta, while NCI-H292 cells were activated for 6 and 14 hours with the following cytokines: 5ng/ml IL-4, 5ng/ml IL-9, 5ng/ml IL-13 and 25ng/ml IFN gamma.

For these cell lines and blood cells, RNA was prepared by lysing approximately 10^7cells/ml using Trizol (Gibco BRL). Briefly, 1/10 volume of bromochloropropane (Molecular Research Corporation) was added to the RNA sample, vortexed and after 10 minutes at room temperature, the tubes were spun at 14,000 rpm in a Sorvall SS34 rotor. The aqueous phase was removed and placed in a 15ml Falcon Tube. An equal volume of isopropanol was added and left at -20°C overnight. The precipitated RNA was spun down at 9,000 rpm for 15 min in a Sorvall SS34 rotor and washed in 70% ethanol. The pellet was redissolved in 300µl of RNAse-free water and 35µl buffer (Promega) 5µl DTT, 7µl RNAsin and 8µl DNAse were added. The tube was incubated at 37°C for 30 minutes to remove contaminating genomic DNA, extracted once with phenol chloroform and re-precipitated with 1/10 volume of 3M sodium acetate and 2 volumes of 100% ethanol. The RNA was spun down and placed in RNAse free water. RNA was stored at -80°C.

AI_comprehensive panel_v1.0

The plates for AI_comprehensive panel_v1.0 include two control wells and 89 test samples comprised of cDNA isolated from surgical and postmortem human tissues obtained from the Backus Hospital and Clinomics (Frederick, MD). Total RNA was extracted from tissue samples from the Backus Hospital in the Facility at CuraGen. Total RNA from other tissues was obtained from Clinomics.

Joint tissues including synovial fluid, synovium, bone and cartilage were obtained from patients undergoing total knee or hip replacement surgery at the Backus Hospital. Tissue samples were immediately snap frozen in liquid nitrogen to ensure that isolated RNA was of optimal quality and not degraded. Additional samples of osteoarthritis and rheumatoid arthritis joint tissues were obtained from Clinomics. Normal control tissues were supplied by Clinomics and were obtained during autopsy of trauma victims.

Surgical specimens of psoriatic tissues and adjacent matched tissues were provided as total RNA by Clinomics. Two male and two female patients were selected between the ages of 25 and 47. None of the patients were taking prescription drugs at the time samples were isolated.

Surgical specimens of diseased colon from patients with ulcerative colitis and Crohns disease and adjacent matched tissues were obtained from Clinomics. Bowel tissue from three female and three male Crohn's patients between the ages of 41-69 were used. Two patients were not on prescription medication while the others were taking dexamethasone, phenobarbital, or tylenol. Ulcerative colitis tissue was from three male and four female patients. Four of the patients were taking lebvid and two were on phenobarbital.

Total RNA from post mortem lung tissue from trauma victims with no disease or with emphysema, asthma or COPD was purchased from Clinomics. Emphysema patients ranged in age from 40-70 and all were smokers, this age range was chosen to focus on patients with cigarette-linked emphysema and to avoid those patients with alpha-1anti-trypsin deficiencies. Asthma patients ranged in age from 36-75, and excluded smokers to prevent those patients that could also have COPD. COPD patients ranged in age from 35-80 and included both smokers and non-smokers. Most patients were taking corticosteroids, and bronchodilators.

In the labels employed to identify tissues in the AI_comprehensive panel_v1.0 panel, the following abbreviations are used:

AI = Autoimmunity
Syn = Synovial
Normal = No apparent disease
Rep22 /Rep20 = individual patients

RA = Rheumatoid arthritis
Backus = From Backus Hospital
OA = Osteoarthritis
(SS) (BA) (MF) = Individual patients
Adj = Adjacent tissue
Match control = adjacent tissues
-M = Male
-F = Female
COPD = Chronic obstructive pulmonary disease

Panels 5D and 5I

The plates for Panel 5D and 5I include two control wells and a variety of cDNAs isolated from human tissues and cell lines with an emphasis on metabolic diseases. Metabolic tissues were obtained from patients enrolled in the Gestational Diabetes study. Cells were obtained during different stages in the differentiation of adipocytes from human mesenchymal stem cells. Human pancreatic islets were also obtained.

In the Gestational Diabetes study subjects are young (18 - 40 years), otherwise healthy women with and without gestational diabetes undergoing routine (elective) Caesarean section. After delivery of the infant, when the surgical incisions were being repaired/closed, the obstetrician removed a small sample.

Patient 2: Diabetic Hispanic, overweight, not on insulin
Patient 7-9: Nondiabetic Caucasian and obese (BMI>30)
Patient 10: Diabetic Hispanic, overweight, on insulin
Patient 11: Nondiabetic African American and overweight
Patient 12: Diabetic Hispanic on insulin

Adipocyte differentiation was induced in donor progenitor cells obtained from Osirus (a division of Clonetics/BioWhittaker) in triplicate, except for Donor 3U which had only two replicates. Scientists at Clonetics isolated, grew and differentiated human mesenchymal stem cells (HuMSCs) for CuraGen based on the published protocol found in Mark F. Pittenger, et al., Multilineage Potential of Adult Human Mesenchymal Stem Cells Science Apr 2 1999: 143-147. Clonetics provided Trizol lysates or frozen pellets suitable for mRNA isolation and ds cDNA production. A general description of each donor is as follows:

Donor 2 and 3 U: Mesenchymal Stem cells, Undifferentiated Adipose
Donor 2 and 3 AM: Adipose, AdiposeMidway Differentiated
Donor 2 and 3 AD: Adipose, Adipose Differentiated

Human cell lines were generally obtained from ATCC (American Type Culture Collection), NCI or the German tumor cell bank and fall into the following tissue groups: kidney proximal convoluted tubule, uterine smooth muscle cells, small intestine, liver HepG2

cancer cells, heart primary stromal cells, and adrenal cortical adenoma cells. These cells are all cultured under standard recommended conditions and RNA extracted using the standard procedures. All samples were processed at CuraGen to produce single stranded cDNA.

Panel 5I contains all samples previously described with the addition of pancreatic islets from a 58 year old female patient obtained from the Diabetes Research Institute at the University of Miami School of Medicine. Islet tissue was processed to total RNA at an outside source and delivered to CuraGen for addition to panel 5I.

In the labels employed to identify tissues in the 5D and 5I panels, the following abbreviations are used:

GO Adipose = Greater Omentum Adipose
SK = Skeletal Muscle
UT = Uterus
PL = Placenta
AD = Adipose Differentiated
AM = Adipose Midway Differentiated
U = Undifferentiated Stem Cells

Panel CNSD.01

The plates for Panel CNSD.01 include two control wells and 94 test samples comprised of cDNA isolated from postmortem human brain tissue obtained from the Harvard Brain Tissue Resource Center. Brains are removed from calvaria of donors between 4 and 24 hours after death, sectioned by neuroanatomists, and frozen at -80°C in liquid nitrogen vapor. All brains are sectioned and examined by neuropathologists to confirm diagnoses with clear associated neuropathology.

Disease diagnoses are taken from patient records. The panel contains two brains from each of the following diagnoses: Alzheimer's disease, Parkinson's disease, Huntington's disease, Progressive Supernuclear Palsy, Depression, and "Normal controls". Within each of these brains, the following regions are represented: cingulate gyrus, temporal pole, globus palladus, substantia nigra, Brodman Area 4 (primary motor strip), Brodman Area 7 (parietal cortex), Brodman Area 9 (prefrontal cortex), and Brodman area 17 (occipital cortex). Not all brain regions are represented in all cases; e.g., Huntington's disease is characterized in part by neurodegeneration in the globus palladus, thus this region is impossible to obtain from confirmed Huntington's cases. Likewise Parkinson's disease is characterized by degeneration of the substantia nigra making this region more difficult to obtain. Normal control brains were

examined for neuropathology and found to be free of any pathology consistent with neurodegeneration.

In the labels employed to identify tissues in the CNS panel, the following abbreviations are used:

PSP = Progressive supranuclear palsy
Sub Nigra = Substantia nigra
Glob Palladus= Globus palladus
Temp Pole = Temporal pole
Cing Gyr = Cingulate gyrus
BA 4 = Brodman Area 4

Panel CNS_Neurodegeneration_V1.0

The plates for Panel CNS_Neurodegeneration_V1.0 include two control wells and 47 test samples comprised of cDNA isolated from postmortem human brain tissue obtained from the Harvard Brain Tissue Resource Center (McLean Hospital) and the Human Brain and Spinal Fluid Resource Center (VA Greater Los Angeles Healthcare System). Brains are removed from calvaria of donors between 4 and 24 hours after death, sectioned by neuroanatomists, and frozen at -80°C in liquid nitrogen vapor. All brains are sectioned and examined by neuropathologists to confirm diagnoses with clear associated neuropathology.

Disease diagnoses are taken from patient records. The panel contains six brains from Alzheimer's disease (AD) patients, and eight brains from "Normal controls" who showed no evidence of dementia prior to death. The eight normal control brains are divided into two categories: Controls with no dementia and no Alzheimer's like pathology (Controls) and controls with no dementia but evidence of severe Alzheimer's like pathology, (specifically senile plaque load rated as level 3 on a scale of 0-3; 0 = no evidence of plaques, 3 = severe AD senile plaque load). Within each of these brains, the following regions are represented: hippocampus, temporal cortex (Brodman Area 21), parietal cortex (Brodman area 7), and occipital cortex (Brodman area 17). These regions were chosen to encompass all levels of neurodegeneration in AD. The hippocampus is a region of early and severe neuronal loss in AD; the temporal cortex is known to show neurodegeneration in AD after the hippocampus; the parietal cortex shows moderate neuronal death in the late stages of the disease; the occipital cortex is spared in AD and therefore acts as a "control" region within AD patients. Not all brain regions are represented in all cases.

In the labels employed to identify tissues in the CNS_Neurodegeneration_V1.0 panel, the following abbreviations are used:

AD = Alzheimer's disease brain; patient was demented and showed AD-like pathology upon autopsy
Control = Control brains; patient not demented, showing no neuropathology
Control (Path) = Control brains; pateint not demented but showing sever AD-like pathology
SupTemporal Ctx = Superior Temporal Cortex
Inf Temporal Ctx = Inferior Temporal Cortex

NOV1

Expression of gene NOV1 was assessed using the primer-probe sets Ag1865, Ag2813, Ag1357, Ag1377, Ag2027, Ag2029 and Ag2039, described in Tables 4A, 4B, 4C, 4D, 4E, 4F and 4G. Results of the RTQ-PCR runs are shown in Tables 4H, 4I, 4J, 4K and 4L.

Table 4A. Probe Name Ag1865

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-AGGCAAAATCATCAACATCAAC-3'	22	1496	18
Probe	TET-5'-CTCAGCCCACCTCACAGAGACCATCT-3'-TAMRA	26	1532	19
Reverse	5'-TTAGCACACCAGGAGACATCTC-3'	22	1570	20

Table 4B. Probe Name Ag2813

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-AATCATCAACATCAACATTGCA-3'	22	1502	21
Probe	TET-5'-CTCAGCCCACCTCACAGAGACCATCT-3'-TAMRA	26	1532	22
Reverse	5'-GTTAGCACACCAGGAGACATCT-3'	22	1571	23

Table 4C. Probe Name Ag1357

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-ATCAGAACTCCTGCCACTCTTT-3'	22	519	24
Probe	TET-5'-TGGACCTATGACACGCAGCAATTCTT-3'-TAMRA	26	470	25
Reverse	5'-ATGCCAATGACTGAGAAAGTTG-3'	22	448	26

Table 4D. Probe Name Ag1377

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-TATTACTTGGGTGGTCATTCCA-3'	22	1069	27
Probe	TET-5'-CAACCCATGGAAAATTCTGGATTTCG-3'-TAMRA	26	1093	28
Reverse	5'-ATATTCCCAATGTTGCCATTTC-3'	22	1142	29

Table 4E. Probe Name Ag2027

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-AGAAAACCACAGGAAGAATGGT-3'	22	2169	30

Probe	TET-5'-CACTTTGTGGGCTGCCCATAATCTTT-3'-TAMRA	26	2197	31
Reverse	5'-ATAATGATGCCTGAACACCAAA-3'	22	2246	32

Table 4F. Probe Name Ag2029

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-AGGCAAAATCATCAACATCAAC-3'	22	1496	33
Probe	TET-5'-CTCAGCCCACCTCACAGAGACCATCT-3'-TAMRA	26	1532	34
Reverse	5'-TTAGCACACCAGGAGACATCTC-3'	22	1570	35

Table 4G. Probe Name Ag2039

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-GACCTATGACACGCAGCAAT-3'	20	474	36
Probe	TET-5'-TCTTCAACTTTCTCAGTCATTGGCAT-3'-TAMRA	26	448	37
Reverse	5'-GGAAGCCATGATTGCATATG-3'	20	399	38

Table 4H. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag1865, Run 207929096	Rel. Exp.(%) Ag2813, Run 209052394	Tissue Name	Rel. Exp.(%) Ag1865, Run 207929096	Rel. Exp.(%) Ag2813, Run 209052394
AD 1 Hippo	13.4	14.1	Control (Path) 3 Temporal Ctx	9.6	11.9
AD 2 Hippo	31.0	33.4	Control (Path) 4 Temporal Ctx	27.5	35.6
AD 3 Hippo	9.0	12.1	AD 1 Occipital Ctx	17.9	22.5
AD 4 Hippo	12.2	14.0	AD 2 Occipital Ctx (Missing)	0.0	0.0
AD 5 Hippo	90.1	**100.0**	AD 3 Occipital Ctx	10.9	11.6
AD 6 Hippo	55.9	58.6	AD 4 Occipital Ctx	22.1	25.5
Control 2 Hippo	30.8	33.7	AD 5 Occipital Ctx	47.0	49.3
Control 4 Hippo	12.2	19.9	AD 5 Occipital Ctx	32.1	36.6
Control (Path) 3 Hippo	11.3	11.3	Control 1 Occipital Ctx	7.4	6.3
AD 1	21.3	27.4	Control 2	46.0	48.6

Temporal Ctx			Occipital Ctx		
AD 2 Temporal Ctx	32.3	33.2	Control 3 Occipital Ctx	15.7	19.5
AD 3 Temporal Ctx	6.3	10.2	Control 4 Occipital Ctx	13.6	12.8
AD 4 Temporal Ctx	24.1	24.3	Control (Path) 1 Occipital Ctx	96.6	79.6
AD 5 Inf Temporal Ctx	**100.0**	**100.0**	Control (Path) 2 Occipital Ctx	10.9	13.1
AD 5 Sup Temporal Ctx	53.6	55.5	Control (Path) 3 Occipital Ctx	8.5	6.5
AD 6 Inf Temporal Ctx	52.1	55.1	Control (Path) 4 Occipital Ctx	13.4	17.7
AD 6 Sup Temporal Ctx	48.3	53.6	Control 1 Parietal Ctx	12.0	12.6
Control 1 Temporal Ctx	7.7	11.7	Control 2 Parietal Ctx	46.0	57.8
Control 2 Temporal Ctx	31.4	37.9	Control 3 Parietal Ctx	14.2	18.6
Control 3 Temporal Ctx	18.0	20.4	Control (Path) 1 Parietal Ctx	58.2	71.2
Control 3 Temporal Ctx	10.4	12.4	Control (Path) 2 Parietal Ctx	25.9	29.1
Control (Path) 1 Temporal Ctx	51.1	51.1	Control (Path) 3 Parietal Ctx	12.8	11.0
Control (Path) 2 Temporal Ctx	28.7	38.7	Control (Path) 4 Parietal Ctx	33.9	45.7

Table 4I. Panel 1.2

Tissue Name	Rel. Exp.(%) Ag1357, Run 134665778	Rel. Exp.(%) Ag1377, Run 134831870	Tissue Name	Rel. Exp.(%) Ag1357, Run 134665778	Rel. Exp.(%) Ag1377, Run 134831870

Endothelial cells	2.6	3.7	Renal ca. 786-0	33.4	32.3
Heart (Fetal)	2.0	2.8	Renal ca. A498	8.6	8.7
Pancreas	2.8	2.6	Renal ca. RXF 393	2.4	1.6
Pancreatic ca. CAPAN 2	5.4	4.7	Renal ca. ACHN	4.8	3.1
Adrenal gland	15.5	16.7	Renal ca. UO-31	1.8	1.3
Thyroid	8.2	8.0	Renal ca. TK-10	4.2	2.4
Salivary gland	17.4	23.5	Liver	13.5	21.8
Pituitary gland	7.0	3.9	Liver (fetal)	3.4	5.3
Brain (fetal)	5.9	7.1	Liver ca. (hepatoblast) HepG2	16.5	12.2
Brain (whole)	22.5	19.3	Lung	4.0	5.0
Brain (amygdala)	10.5	12.3	Lung (fetal)	2.6	4.7
Brain (cerebellum)	8.7	10.1	Lung ca. (small cell) LX-1	20.7	16.7
Brain (hippocampus)	14.2	24.7	Lung ca. (small cell) NCI-H69	5.1	6.2
Brain (thalamus)	12.4	15.0	Lung ca. (s.cell var.) SHP-77	2.3	1.9
Cerebral Cortex	17.9	24.5	Lung ca. (large cell)NCI-H460	14.4	16.0
Spinal cord	9.0	7.4	Lung ca. (non-sm. cell) A549	9.2	9.3
glio/astro U87-MG	8.7	12.7	Lung ca. (non-s.cell) NCI-H23	3.5	1.9
glio/astro U-118-MG	3.1	2.2	Lung ca. (non-s.cell) HOP-62	4.3	2.9
astrocytoma SW1783	2.4	3.5	Lung ca. (non-s.cl) NCI-H522	13.6	11.8
neuro*; met SK-N-AS	47.6	**100.0**	Lung ca. (squam.) SW 900	8.8	8.0
astrocytoma SF-539	4.7	2.9	Lung ca. (squam.) NCI-H596	1.0	1.6
astrocytoma SNB-75	1.2	0.8	Mammary gland	4.8	4.4
glioma SNB-19	13.2	5.7	Breast ca.* (pl.ef) MCF-7	44.1	39.8

glioma U251	0.9	0.7	Breast ca.* (pl.ef) MDA-MB-231	7.8	14.6
glioma SF-295	1.7	0.9	Breast ca.* (pl. ef) T47D	11.7	22.4
Heart	38.2	48.3	Breast ca. BT-549	12.3	15.8
Skeletal muscle	23.0	43.2	Breast ca. MDA-N	32.8	23.0
Bone marrow	2.1	3.3	Ovary	1.1	0.9
Thymus	2.9	2.6	Ovarian ca. OVCAR-3	7.2	4.3
Spleen	3.8	4.9	Ovarian ca. OVCAR-4	7.4	4.7
Lymph node	2.5	3.7	Ovarian ca. OVCAR-5	18.9	13.4
Colorectal	0.8	1.6	Ovarian ca. OVCAR-8	9.3	20.0
Stomach	7.4	6.3	Ovarian ca. IGROV-1	14.3	6.9
Small intestine	16.6	22.2	Ovarian ca. (ascites) SK-OV-3	8.7	5.6
Colon ca. SW480	0.8	0.7	Uterus	7.2	6.9
Colon ca.* SW620 (SW480 met)	4.2	8.5	Placenta	14.7	30.1
Colon ca. HT29	14.1	14.4	Prostate	6.2	10.3
Colon ca. HCT-116	17.0	6.6	Prostate ca.* (bone met) PC-3	12.2	14.3
Colon ca. CaCo-2	21.5	21.6	Testis	5.9	4.1
CC Well to Mod Diff (ODO3866)	3.2	3.2	Melanoma Hs688(A).T	1.6	1.5
Colon ca. HCC-2998	**100.0**	42.9	Melanoma* (met) Hs688(B).T	0.6	1.0
Gastric ca. (liver met) NCI-N87	25.0	31.2	Melanoma UACC-62	1.8	1.4
Bladder	7.6	9.3	Melanoma M14	4.1	2.1
Trachea	3.7	4.0	Melanoma LOX IMVI	0.9	0.8
Kidney	18.7	25.2	Melanoma*	4.9	5.6

			(met) SK-MEL-5		
Kidney (fetal)	8.4	9.3			

Table 4J. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag1865, Run 148610768	Rel. Exp.(%) Ag2027, Run 165549729	Rel. Exp.(%) Ag2029, Run 165552311	Rel. Exp.(%) Ag2039, Run 146127643	Rel. Exp.(%) Ag2813, Run 165528059
Liver adenocarcinoma	10.2	9.6	4.3	13.5	9.1
Pancreas	5.0	4.8	8.3	6.0	7.0
Pancreatic ca. CAPAN 2	8.9	10.8	20.3	21.5	24.8
Adrenal gland	13.2	12.9	17.8	13.2	17.8
Thyroid	9.7	17.2	15.1	14.0	11.2
Salivary gland	3.9	13.7	8.1	6.0	11.3
Pituitary gland	7.6	8.4	9.6	10.9	8.8
Brain (fetal)	10.2	28.7	34.9	12.4	36.3
Brain (whole)	43.5	88.9	**100.0**	40.9	**100.0**
Brain (amygdala)	36.1	65.1	61.1	22.1	44.8
Brain (cerebellum)	10.0	41.2	54.3	10.4	38.2
Brain (hippocampus)	**100.0**	47.3	57.4	52.9	63.7
Brain (substantia nigra)	12.4	48.3	43.2	13.0	45.4
Brain (thalamus)	34.4	**100.0**	72.2	28.9	81.8
Cerebral Cortex	50.0	4.8	13.6	63.7	11.6
Spinal cord	19.6	55.5	68.8	35.1	44.1
glio/astro U87-MG	6.7	11.9	13.5	27.5	10.7
glio/astro U-118-MG	13.5	24.1	32.8	26.8	19.9
astrocytoma SW1783	8.7	9.5	13.2	25.2	11.2
neuro*; met SK-N-AS	52.9	35.6	56.6	**100.0**	51.8
astrocytoma SF-539	9.6	8.0	11.8	21.3	14.5
astrocytoma SNB-75	21.8	21.6	22.8	54.3	20.3
glioma SNB-19	11.2	17.0	17.1	24.5	22.7
glioma U251	3.2	11.4	19.8	6.8	14.3
glioma SF-295	10.3	5.8	6.6	9.6	8.7
Heart (Fetal)	6.4	0.0	1.0	9.0	0.7

Heart	7.5	35.1	30.4	12.9	34.4
Skeletal muscle (Fetal)	19.9	0.0	0.0	25.5	0.4
Skeletal muscle	6.7	89.5	73.7	11.1	52.5
Bone marrow	4.0	6.3	6.3	5.1	5.4
Thymus	4.5	2.9	5.1	8.4	6.4
Spleen	14.7	11.7	12.6	14.1	11.6
Lymph node	9.2	25.9	23.3	9.2	21.8
Colorectal	8.4	1.5	3.1	7.4	3.9
Stomach	20.0	12.2	17.3	15.4	11.4
Small intestine	12.8	33.0	31.4	12.5	18.9
Colon ca. SW480	10.0	3.1	3.9	17.3	4.3
Colon ca.* SW620 (SW480 met)	8.3	4.7	7.7	14.2	9.7
Colon ca. HT29	12.3	4.1	6.6	22.8	9.6
Colon ca. HCT-116	6.8	3.2	6.7	17.9	9.0
Colon ca. CaCo-2	25.3	13.0	16.4	49.7	17.4
CC Well to Mod Diff (ODO3866)	8.1	11.0	16.0	0.0	14.6
Colon ca. HCC-2998	20.9	9.3	18.9	43.5	19.6
Gastric ca. (liver met) NCI-N87	20.0	17.9	31.2	68.8	30.8
Bladder	11.3	9.5	17.7	13.9	11.8
Trachea	13.2	12.5	10.9	18.0	11.0
Kidney	7.9	22.2	23.8	7.5	18.2
Kidney (fetal)	6.7	13.1	13.5	9.5	14.8
Renal ca. 786-0	24.0	23.3	30.4	44.8	28.9
Renal ca. A498	24.5	21.3	37.9	48.0	36.9
Renal ca. RXF 393	5.2	11.3	15.4	6.9	19.6
Renal ca. ACHN	10.6	7.9	13.7	26.8	12.7
Renal ca. UO-31	12.3	11.0	20.2	15.3	16.8
Renal ca. TK-10	3.2	8.0	12.1	10.4	8.8
Liver	10.7	37.1	23.0	9.9	30.1
Liver (fetal)	6.9	11.3	19.3	6.5	16.4
Liver ca. (hepatoblast) HepG2	13.0	20.9	34.2	39.0	33.0
Lung	9.0	17.2	12.9	11.3	11.9
Lung (fetal)	12.6	16.6	13.6	18.4	19.6
Lung ca. (small	14.6	18.3	29.9	33.4	29.7

cell) LX-1					
Lung ca. (small cell) NCI-H69	10.4	3.3	6.0	19.1	7.8
Lung ca. (s.cell var.) SHP-77	11.5	12.1	23.8	31.4	15.9
Lung ca. (large cell)NCI-H460	5.2	13.1	25.2	8.1	21.8
Lung ca. (non-sm. cell) A549	12.9	7.2	8.0	23.5	14.6
Lung ca. (non-s.cell) NCI-H23	14.9	8.4	8.9	21.3	5.3
Lung ca. (non-s.cell) HOP-62	4.3	4.9	5.1	6.9	8.2
Lung ca. (non-s.cl) NCI-H522	24.1	4.9	11.0	23.2	14.8
Lung ca. (squam.) SW 900	9.0	12.7	19.9	28.1	20.7
Lung ca. (squam.) NCI-H596	3.0	1.8	6.3	6.6	6.9
Mammary gland	11.2	8.7	14.2	24.0	9.6
Breast ca.* (pl.ef) MCF-7	13.5	21.6	34.6	44.4	21.9
Breast ca.* (pl.ef) MDA-MB-231	15.0	22.8	46.7	51.8	38.2
Breast ca.* (pl. ef) T47D	7.0	6.2	8.7	14.3	6.7
Breast ca. BT-549	18.0	24.3	34.9	38.2	28.3
Breast ca. MDA-N	13.1	5.9	8.3	49.0	11.4
Ovary	9.0	1.1	0.6	6.7	0.8
Ovarian ca. OVCAR-3	6.3	9.7	17.2	22.4	20.9
Ovarian ca. OVCAR-4	2.8	7.1	17.6	4.7	13.7
Ovarian ca. OVCAR-5	20.3	13.4	23.2	27.0	21.3
Ovarian ca. OVCAR-8	11.7	5.1	7.4	17.1	9.1
Ovarian ca. IGROV-1	7.6	4.3	6.7	14.3	7.8
Ovarian ca. (ascites) SK-OV-3	15.8	15.6	24.1	0.0	37.4
Uterus	5.6	17.8	20.9	7.7	17.4
Placenta	6.2	7.6	14.9	12.7	8.4

Prostate	7.4	6.4	13.3	8.8	5.8
Prostate ca.* (bone met) PC-3	8.2	6.3	7.9	13.5	9.9
Testis	58.2	25.2	18.9	38.4	36.3
Melanoma Hs688(A).T	7.4	3.2	4.7	35.4	4.2
Melanoma* (met) Hs688(B).T	8.1	4.5	4.2	36.3	4.0
Melanoma UACC-62	0.4	2.1	2.5	1.1	2.9
Melanoma M14	9.5	22.2	32.3	10.7	34.6
Melanoma LOX IMVI	2.3	0.5	1.4	4.3	2.0
Melanoma* (met) SK-MEL-5	7.1	6.7	8.8	15.1	12.2
Adipose	8.6	7.6	9.3	10.7	8.0

Table 4K. Panel 2D

Tissue Name	Rel. Exp.(%) Ag2813, Run 162599032	Tissue Name	Rel. Exp.(%) Ag2813, Run 162599032
Normal Colon	56.6	Kidney Margin 8120608	4.5
CC Well to Mod Diff (ODO3866)	18.8	Kidney Cancer 8120613	11.2
CC Margin (ODO3866)	14.4	Kidney Margin 8120614	6.8
CC Gr.2 rectosigmoid (ODO3868)	13.4	Kidney Cancer 9010320	6.8
CC Margin (ODO3868)	3.4	Kidney Margin 9010321	11.9
CC Mod Diff (ODO3920)	10.9	Normal Uterus	3.5
CC Margin (ODO3920)	9.5	Uterine Cancer 064011	18.8
CC Gr.2 ascend colon (ODO3921)	28.1	Normal Thyroid	21.3
CC Margin (ODO3921)	9.8	Thyroid Cancer	16.8
CC from Partial Hepatectomy (ODO4309) Mets	60.7	Thyroid Cancer A302152	20.3
Liver Margin (ODO4309)	56.6	Thyroid Margin A302153	21.8
Colon mets to lung (OD04451-01)	8.7	Normal Breast	16.6
Lung Margin (OD04451-02)	8.1	Breast Cancer	13.1

Normal Prostate 6546-1	**100.0**	Breast Cancer (OD04590-01)	51.1
Prostate Cancer (OD04410)	31.6	Breast Cancer Mets (OD04590-03)	54.3
Prostate Margin (OD04410)	27.0	Breast Cancer Metastasis	37.9
Prostate Cancer (OD04720-01)	10.0	Breast Cancer	11.3
Prostate Margin (OD04720-02)	19.5	Breast Cancer	13.7
Normal Lung	35.6	Breast Cancer 9100266	24.0
Lung Met to Muscle (ODO4286)	20.4	Breast Margin 9100265	6.8
Muscle Margin (ODO4286)	8.6	Breast Cancer A209073	17.0
Lung Malignant Cancer (OD03126)	29.7	Breast Margin A2090734	8.4
Lung Margin (OD03126)	26.6	Normal Liver	36.9
Lung Cancer (OD04404)	17.9	Liver Cancer	19.2
Lung Margin (OD04404)	10.9	Liver Cancer 1025	15.7
Lung Cancer (OD04565)	4.7	Liver Cancer 1026	4.8
Lung Margin (OD04565)	5.0	Liver Cancer 6004-T	18.0
Lung Cancer (OD04237-01)	32.1	Liver Tissue 6004-N	9.1
Lung Margin (OD04237-02)	19.5	Liver Cancer 6005-T	7.6
Ocular Mel Met to Liver (ODO4310)	9.8	Liver Tissue 6005-N	4.9
Liver Margin (ODO4310)	44.8	Normal Bladder	25.9
Melanoma Metastasis	9.4	Bladder Cancer	4.6
Lung Margin (OD04321)	18.6	Bladder Cancer	11.1
Normal Kidney	69.3	Bladder Cancer (OD04718-01)	31.4
Kidney Ca, Nuclear grade 2 (OD04338)	58.6	Bladder Normal Adjacent (OD04718-03)	9.5
Kidney Margin (OD04338)	24.7	Normal Ovary	1.7
Kidney Ca Nuclear grade 1/2 (OD04339)	17.3	Ovarian Cancer	18.7
Kidney Margin (OD04339)	58.2	Ovarian Cancer (OD04768-07)	65.5
Kidney Ca, Clear cell type (OD04340)	36.6	Ovary Margin (OD04768-08)	4.4

Kidney Margin (OD04340)	33.9	Normal Stomach	12.9
Kidney Ca, Nuclear grade 3 (OD04348)	9.8	Gastric Cancer 9060358	3.4
Kidney Margin (OD04348)	29.3	Stomach Margin 9060359	12.2
Kidney Cancer (OD04622-01)	19.1	Gastric Cancer 9060395	13.8
Kidney Margin (OD04622-03)	4.6	Stomach Margin 9060394	9.8
Kidney Cancer (OD04450-01)	49.7	Gastric Cancer 9060397	36.1
Kidney Margin (OD04450-03)	44.4	Stomach Margin 9060396	2.6
Kidney Cancer 8120607	3.1	Gastric Cancer 064005	70.7

Table 4L. Panel 4D

Tissue Name	Rel. Exp.(%) Ag1377, Run 145703126	Rel. Exp.(%) Ag1865, Run 164038695	Rel. Exp.(%) Ag2027, Run 160997517	Rel. Exp.(%) Ag2029, Run 161379206	Rel. Exp.(%) Ag2039, Run 156265871	Rel. Exp.(%) Ag2813, Run 162293678
Secondary Th1 act	20.2	22.4	18.0	15.8	24.7	17.2
Secondary Th2 act	14.7	25.7	24.5	17.6	18.7	26.8
Secondary Tr1 act	21.2	15.5	21.3	20.7	21.9	18.9
Secondary Th1 rest	7.9	10.2	8.4	4.5	9.2	8.4
Secondary Th2 rest	12.6	8.6	12.2	9.3	7.1	10.8
Secondary Tr1 rest	8.0	14.8	17.4	9.8	10.2	13.8
Primary Th1 act	22.8	20.3	18.8	16.0	15.0	18.4
Primary Th2 act	23.3	28.1	17.1	15.6	18.6	18.4
Primary Tr1 act	30.8	24.1	24.7	17.2	30.1	26.4
Primary Th1 rest	46.7	49.0	48.3	33.7	65.5	49.0
Primary Th2 rest	19.2	17.7	29.1	18.6	29.7	23.8
Primary Tr1 rest	14.1	18.2	20.2	13.8	20.3	22.2
CD45RA CD4 lymphocyte act	12.1	17.8	12.2	9.3	13.2	13.5
CD45RO CD4 lymphocyte act	20.7	33.9	16.5	14.3	21.3	24.0
CD8 lymphocyte act	16.6	17.2	14.7	13.3	12.2	18.2
Secondary CD8 lymphocyte rest	18.3	22.4	19.8	16.2	13.1	20.0
Secondary CD8 lymphocyte act	14.4	17.4	13.7	12.1	14.0	13.1

CD4 lymphocyte none	6.5	9.2	7.0	6.8	7.5	5.7
2ry Th1/Th2/Tr1_anti-CD95 CH11	19.5	17.1	13.1	13.0	16.4	20.2
LAK cells rest	35.1	33.7	31.6	24.8	27.9	37.1
LAK cells IL-2	30.6	34.6	28.7	22.7	20.6	30.6
LAK cells IL-2+IL-12	21.3	21.3	17.4	15.2	14.6	17.8
LAK cells IL-2+IFN gamma	31.4	36.9	27.7	23.7	30.1	34.4
LAK cells IL-2+ IL-18	20.9	21.8	30.8	17.9	33.2	26.2
LAK cells PMA/ionomycin	9.4	13.5	9.5	11.0	15.7	13.9
NK Cells IL-2 rest	11.7	32.5	19.3	17.7	22.5	28.9
Two Way MLR 3 day	19.9	56.3	24.7	21.9	24.0	26.1
Two Way MLR 5 day	12.4	18.7	16.8	14.3	14.4	17.9
Two Way MLR 7 day	8.6	13.4	9.0	6.9	10.5	10.6
PBMC rest	9.9	18.0	9.2	9.5	12.2	12.9
PBMC PWM	70.7	85.9	48.3	53.6	58.2	74.7
PBMC PHA-L	35.4	31.9	18.3	14.5	23.5	24.5
Ramos (B cell) none	27.5	21.6	22.5	16.8	17.9	23.3
Ramos (B cell) ionomycin	72.2	97.3	88.9	80.7	97.3	**100.0**
B lymphocytes PWM	57.8	**100.0**	**100.0**	**100.0**	**100.0**	97.9
B lymphocytes CD40L and IL-4	15.0	23.5	25.0	20.0	24.8	28.9
EOL-1 dbcAMP	7.0	18.2	13.4	12.1	12.2	15.2
EOL-1 dbcAMP PMA/ionomycin	9.0	19.1	23.7	14.9	14.1	18.2
Dendritic cells none	19.6	35.8	24.7	21.3	21.2	29.7
Dendritic cells LPS	15.5	36.3	41.8	36.3	39.5	43.8
Dendritic cells anti-CD40	13.9	39.5	32.3	24.3	25.5	21.6
Monocytes rest	21.2	28.7	31.2	19.9	19.1	20.3
Monocytes LPS	22.1	30.1	20.9	24.0	21.2	21.9
Macrophages rest	26.4	38.2	40.1	30.8	30.6	34.4
Macrophages LPS	61.1	44.4	29.5	29.5	41.5	31.4

HUVEC none	11.2	17.9	24.0	17.6	18.3	19.9
HUVEC starved	26.2	38.7	34.6	19.9	36.1	34.6
HUVEC IL-1beta	4.9	12.9	8.9	7.3	11.7	7.1
HUVEC IFN gamma	14.2	27.9	21.5	17.6	17.8	24.5
HUVEC TNF alpha + IFN gamma	6.5	28.5	20.4	21.0	18.3	27.0
HUVEC TNF alpha + IL4	8.8	19.2	17.1	14.7	63.7	14.9
HUVEC IL-11	7.1	11.1	9.2	7.4	8.8	7.6
Lung Microvascular EC none	13.6	32.1	27.5	21.5	15.0	26.8
Lung Microvascular EC TNFalpha + IL-1beta	11.0	28.3	28.5	20.9	17.4	24.8
Microvascular Dermal EC none	34.4	35.4	35.4	31.6	29.5	38.7
Microsvasular Dermal EC TNFalpha + IL-1beta	15.7	25.0	23.7	18.9	21.3	23.5
Bronchial epithelium TNFalpha + IL1beta	11.4	20.0	2.2	2.0	0.7	18.7
Small airway epithelium none	3.6	6.8	4.2	6.1	6.4	6.9
Small airway epithelium TNFalpha + IL-1beta	26.6	40.3	36.3	30.1	44.1	35.4
Coronery artery SMC rest	13.1	17.7	16.5	13.8	20.9	21.2
Coronery artery SMC TNFalpha + IL-1beta	5.6	13.4	9.5	10.7	8.7	10.0
Astrocytes rest	4.7	18.6	21.2	13.7	12.9	16.7
Astrocytes TNFalpha + IL-1beta	8.0	15.1	10.9	10.8	8.3	11.7
KU-812 (Basophil) rest	6.8	9.9	7.1	7.4	5.4	9.3
KU-812 (Basophil) PMA/ionomycin	20.6	31.4	21.9	21.0	24.0	26.1

CCD1106 (Keratinocytes) none	6.8	17.3	18.7	8.8	12.7	17.0
CCD1106 (Keratinocytes) TNFalpha + IL-1beta	3.6	7.9	1.1	1.2	1.9	7.0
Liver cirrhosis	8.7	8.0	7.0	4.8	4.9	8.0
Lupus kidney	15.1	4.8	4.3	4.0	3.3	5.0
NCI-H292 none	27.2	39.2	35.4	37.4	34.9	40.3
NCI-H292 IL-4	34.2	54.0	55.5	43.5	75.3	54.0
NCI-H292 IL-9	30.8	52.9	70.2	48.0	59.0	57.0
NCI-H292 IL-13	18.0	40.6	32.3	28.9	31.2	16.3
NCI-H292 IFN gamma	13.6	49.3	39.0	39.8	52.1	48.0
HPAEC none	10.9	35.8	13.4	11.3	9.3	10.2
HPAEC TNF alpha + IL-1 beta	10.7	20.0	11.8	11.4	19.3	16.7
Lung fibroblast none	14.6	14.6	7.5	12.1	9.7	12.2
Lung fibroblast TNF alpha + IL-1 beta	9.1	21.0	10.7	12.7	15.1	15.8
Lung fibroblast IL-4	11.6	24.0	22.7	19.9	15.5	23.5
Lung fibroblast IL-9	9.9	24.7	20.9	19.3	14.5	17.9
Lung fibroblast IL-13	21.5	14.2	11.7	9.7	11.0	12.9
Lung fibroblast IFN gamma	22.8	42.6	46.0	39.0	37.9	37.9
Dermal fibroblast CCD1070 rest	20.9	37.6	36.6	31.6	32.3	34.2
Dermal fibroblast CCD1070 TNF alpha	39.5	70.2	89.5	60.3	63.7	65.1
Dermal fibroblast CCD1070 IL-1 beta	11.2	26.2	24.1	18.8	20.2	18.6
Dermal fibroblast IFN gamma	9.5	23.5	19.5	19.8	19.9	18.6
Dermal fibroblast IL-4	14.5	24.8	20.6	16.4	21.5	23.0
IBD Colitis 2	3.3	1.6	1.2	0.5	0.8	1.8
IBD Crohn's	3.4	4.5	4.2	2.7	2.9	4.2
Colon	26.6	52.5	82.4	41.8	51.8	53.2

Lung	10.9	24.1	21.3	17.1	20.3	23.5
Thymus	**100.0**	80.7	76.3	57.4	72.7	76.3
Kidney	19.9	28.7	27.9	25.5	23.8	27.5

CNS_neurodegeneration_v1.0 Summary: Ag1865/Ag2813

Results from two experiments using different probe/primer sets are in good agreement. A NOV1 gene is moderately expressed in all samples in this panel, confirming the expression of this gene in the brain. Gene expression levels are similar in all brain regions tested (hippocampus, temporal cortex, parietal cortex, and occipital cortex) and show no apparent alteration in patients with Alzheimer's disease.

Panel 1.2 Summary: Ag1357/Ag1377

A NOV1 gene is expressed at moderate to high levels in all of the tissues on this panel, with highest expression detected in a brain cancer cell line and a colon cancer cell line (CTs=23.5). In general, expression of this gene appears to be higher in cancer cell lines than in normal tissues; this pattern of expression holds true for breast and ovarian cancer cell lines as well. Thus, expression of this gene may be useful for identifying colon or brain cancer. Furthermore, therapeutic modulation of the expression or function of this gene, through the use of small molecule drugs, antibodies or protein therapeutics, might be of benefit for the treatment of cancer.

Among tissues derived from the central nervous system, a NOV1 gene is expressed at high levels in fetal brain, amygdala, cerebellum, hippocampus, thalamus, cerebral cortex and spinal cord (CTs=25-27). High expression throughout the brain indicates a potential role for this gene in normal brain function. A NOV1 gene encodes a protein with homology to endozepine. Endogenous benzodiazepine-like substances are thought to play a role in the development of hepatic encephalopathy . Gooday R, Hayes PC, Bzeizi K, O'Carroll RE. Benzodiazepine receptor antagonism improves reaction time in latent hepatic encephalopathy. Psychopharmacology (Berl) 1995 Jun;119(3):295-8. It has been suggested that benzodiazepine receptor antagonism may improve cognitive function, particularly speed of information processing, in patients with latent hepatic encephalopathy. Thus, a NOV1 gene product used as a protein therapeutic, or drugs that stimulate the protein's function, may have efficacy in the treatment of hepatic encephalopathy. Increased expression of diazepam binding inhibitor (DBI), a endozepine peptide with anxiogenic action, in Alzheimer's disease, addiction and schizophrenia, indicates that drugs that inhibit NOV1 protein activity may also have utility in

the treatment of these diseases. Edgar PF, Schonberger SJ, Dean B, Faull RL, Kydd R, Cooper GJ. A comparative proteome analysis of hippocampal tissue from schizophrenic and Alzheimer's disease individuals. Mol Psychiatry 1999 Mar;4(2):173-8. Ohkuma S, Katsura M, Tsujimura A. Alterations in cerebral diazepam binding inhibitor expression in drug dependence: a possible biochemical alteration common to drug dependence. Life Sci 2001 Feb 2;68(11):1215-22.

Among tissues with metabolic function, this gene is expressed at high levels in thyroid, adrenal gland, pituitary gland, heart, skeletal muscle, and at moderate levels in pancreas. Therefore, this gene product may be important for the pathogenesis and treatment of diseases in any or all of these tissues. In particular, a diazepam binding inhibitor has previously been identified as a potential autoantigen in autoimmune diabetes in a screen of a human pancreatic islet cDNA library with the sera from the diabetic patients autoantigens in autoimmune diabetes. Suk K, Kim YH, Hwang DY, Ihm SH, Yoo HJ, Lee MS. Molecular cloning and expression of a novel human cDNA related to the diazepam binding inhibitor. Biochim Biophys Acta 1999 May 31;1454(1):126-31.

In addition, treatment of rodents with the octadecaneuropeptide [diazepam-binding inhibitor (33-50)] has been shown to result in decreased food intake and weight loss (ref. 5). These observations suggest that therapeutic modulation of the diazepam-binding inhibitor-like protein encoded by a NOV1 gene may be useful in the treatment of diabetes and obesity.

References:

Endogenous benzodiazepine-like substances are thought to play a role in the development of hepatic encephalopathy (HE). Ten patients with sub-clinical or latent hepatic encephalopathy (LHE) and ten normal controls were cognitively assessed pre- and post-infusion of 0.2 mg of the benzodiazepine (BZ) antagonist flumazenil in a placebo-controlled, cross-over, double-blind design. Flumazenil infusion resulted in a significant improvement in simple reaction time in patients, but not in controls. Saline infusion had no effect on any of the cognitive measures in either group. Flumazenil appeared to have a particular enhancing effect on the cognitive, as opposed to the motor, component of the reaction time task. This finding supports the view that the benzodiazepine/GABA system is implicated in the bradyphrenia that is characteristic of chronic liver disease, even before hepatic encephalopathy is apparent. We conclude that benzodiazepine receptor antagonism may improve cognitive function, particularly speed of information processing, in patients with latent hepatic encephalopathy.

PMID: 7675964

2. Edgar PF, Schonberger SJ, Dean B, Faull RL, Kydd R, Cooper GJ. A comparative proteome analysis of hippocampal tissue from schizophrenic and Alzheimer's disease individuals. Mol Psychiatry 1999 Mar;4(2):173-8.

The proteins expressed by a genome have been termed the proteome. Comparative proteome analysis of brain tissue offers a novel means to identify biologically significant gene products that underlie psychopathology. In this study we collected post mortem hippocampal tissue from the brains of seven schizophrenic, seven Alzheimer's disease (AD) and seven control individuals. Hippocampal proteomes were visualised by two-dimensional gel electrophoresis of homogenised tissue. A mean of 549 (s.d. 35) proteins were successfully matched between each disease group and the control group. In comparison with the control hippocampal proteome, eight proteins in the schizophrenic hippocampal proteome were found to be decreased and eight increased in concentration, whereas, in the AD hippocampal proteome, 35 proteins were decreased and 73 were increased in concentration (P

PMID: 10208449

Mechanisms for formation of drug dependence and expression of withdrawal syndrome have not fully clarified despite of huge accumulation of experimental and clinical data at present. Several clinical features of withdrawal syndrome are considered to be common among patients with drug dependence induced by different drugs of abuse. One of them is anxiety. Recent investigations have revealed that diazepam binding inhibitor (DBI), a peptide consisting of 87 amino acids with molecular weight of about 10 kDa, serves as an inverse agonist for benzodiazepine (BZD) receptors with endogenously anxiogenic potential. These lines of data suggest that cerebral DBI expression in brain may participates in formation of drug dependence and emergence of withdrawal syndrome. Based on this working hypothesis, we have examined DBI expression in the brain derived from mice depended on alcohol (ethanol), nicotine, and morphine to investigate functional relationship between cerebral DBI expression and drug dependence. Cerebral DBI expression significantly increases in animals with drug dependence induced by these drugs, and in the cases of nicotine- and morphine-dependent mice concomitant administration of antagonists for nicotinic acetylcholine and opioid receptors, respectively, abolished the increase. Abrupt cessation of administration of drugs facilitated further increase in DBI expression. Therefore, these alterations in DBI expression have close relationship with formation of drug dependence and emergence of withdrawal syndrome, and are considered to be a common biochemical process in drug dependence induced by different drugs of abuse.

Finding and elucidation of mechanisms for common biochemical alterations among drug dependence may provide a clue to clarify mechanisms for formation of drug dependence and emergence of withdrawal syndrome.

PMID: 11233989

4. Suk K, Kim YH, Hwang DY, Ihm SH, Yoo HJ, Lee MS. Molecular cloning and expression of a novel human cDNA related to the diazepam binding inhibitor. Biochim Biophys Acta 1999 May 31;1454(1):126-31.

In order to isolate the unidentified autoantigens in autoimmune diabetes, a human pancreatic islet cDNA library was constructed and screened with the sera from the diabetic patients. From the library screening, one clone (DRS-1) that strongly reacted with the sera was isolated. Subsequent sequence analysis revealed that the clone was a novel cDNA related to the diazepam binding inhibitor. DRS-1 was expressed in most tissues including liver, lung, tonsil, and thymus, in addition to pancreatic islets. DRS-1 was in vitro translated and the recombinant DRS-1 protein was expressed in Escherichia coli and purified. The size of the in vitro translated or bacterially expressed DRS-1 protein was in agreement with the conceptually translated polypeptide of DRS-1 cDNA. Further studies are required to test whether or not DRS-1 is a new autoantigen in autoimmune diabetes.

PMID: 10354522

The effects of intracerebroventricular administration of the octadecaneuropeptide ODN on food intake have been investigated in rat and mouse. In rats deprived of food from 9:00 a.m. to 7:00 p.m., i.c.v. injection of ODN (30 to 100 ng) provoked a dose-dependent reduction of food consumption during the following 12-h nocturnal period. At a dose of 100 ng, ODN almost completely suppressed food intake. Treatment of rats with diazepam (2 mg/kg s.c.; 15 min before ODN administration) did not affect the anorexigenic response evoked by 100 ng ODN. Continuous i.c.v. infusion of ODN (10 ng/h during 15 days) using osmotic minipumps, significantly reduced food intake during the 2nd, 3rd and 4th days of treatment. The decrease in food consumption was associated with a significant reduction in body weight, which persisted during the 15-day duration of the experiment. In mice deprived of food for 18 h, i.c.v. administration of a low dose of ODN (5 ng) significantly reduced food intake. Treatment of mice with diazepam (1 mg/kg s.c.; 10 min before ODN administration) did not prevent the inhibitory effect of ODN (100 ng) on food intake. The C-terminal octapeptide fragment of ODN mimicked the anorexigenic effect of the intact peptide. Taken together, the present data demonstrate that i.c.v. injection of ODN causes, in both rat and mouse, a long-lasting

anorexigenic effect that is not mediated through central-type benzodiazepine receptors. The biologically active region of ODN appears to be located in the C-terminal domain of the peptide.

PMID: 11239923

Panel 1.3D Summary: Ag1865/2027/2029/2039/2813

Results from multiple experiments with different probes and primer sets show highest expression of a NOV1 gene in regions of the brain. The expression of this gene is highest in a samples derived from whole brain tissue (CT=27). High levels of expression are detected in whole brain, fetal brain, amygdala, cerebellum, hippocampus, thalamus, cerebral cortex and spinal cord (CTs=28-31). Please see Panel 1.2 summary for discussion of potential utility based upon expression in the CNS.

Interestingly, a NOV1 gene is more highly expressed in fetal skeletal muscle (CT=29) than adult skeletal muscle (CT=36-40), suggesting that this gene could be used to distinguish the two. In addition, the increased 20936375_0_1_da1 gene expression in fetal skeletal muscle suggests that the protein product may enhance muscular growth or development in the fetus and thus may also act in a regenerative capacity in the adult. Therefore, therapeutic modulation of the 20936375_0_1_da1 gene could be useful in treatment of muscular related disease. More specifically, treatment of weak or dystrophic muscle with the protein encoded by this gene could restore muscle mass or function.

Among tissues with metabolic function, a NOV1 gene is expressed at low to moderate levels in pancreas, adrenal gland, pituitary gland, thyroid, heart, liver and adipose. Please see Panel 1.2 summary for discussion of potential utility of this gene in metabolic diseases.

Across all of the panels, there appears to be consistent expression of a NOV1 gene in the neuroblastoma cell line. This suggests that the expression of this gene could be used to distinguish this cell type from others. Moreover, therapeutic modulation of this gene, through the use of small molecule drugs, antibodies or protein therapeutics might be of use in the treatment of neuroblastoma.

Panel 2D Summary: Ag2813

A NOV1 gene is most highly expressed in a sample derived from normal prostate tissue (CT=25.6). In addition, substantial expression of this gene is detected in a number of tissue samples on this panel. Strikingly, NOV1 gene expression is higher in cancers of the ovary, breast and stomach when compared with their associated normal adjacent tissues. Thus, the

expression of this gene could be used to distinguish ovarian, breast or stomach cancer tissue from normal tissue. Moreover, therapeutic modulation of a NOV1 gene expression or activity, through the use of small molecule drugs, antibodies or protein therapeutics might be of benefit in the treatment of ovarian, breast or stomach cancer.

Panel 4D Summary: Ag1377/Ag1865/Ag2027/Ag2029/Ag2039/Ag2813

Multiple experiments using different probe/primer sets gave results that are in very good agreement. A NOV1 gene is moderately expressed in the majority of samples on this panel (CT values ranging from 27.5 to 32). However, this gene is expressed at a higher level in activated B cells (Ramos, B cells plus PWM and PBMC plus PWM). A NOV1 gene encodes a protein with homology to a membrane-associated diazepam binding inhibitor, which has been shown to have immunomodulatory activity. Stepien H, Agro A, Crossley J, Padol I, Richards C, Stanisz A. Immunomodulatory properties of diazepam-binding inhibitor: effect on human interleukin-6 secretion, lymphocyte proliferation and natural killer cell activity in vitro. Neuropeptides 1993 Sep;25(3):207-11

Therefore small molecules target or antibodies against a NOV1 protein might modulate B cell activity and be useful in the treatment of diseases associated with B cell activation, such as autoiommune diseases (including systemic lupus erythematosus and rheumatoid arthritis) and hyperglobulinemia. In addition, a NOV1 gene is quite abundantly expressed in dermal fibroblasts treated with TNF-alpha, suggesting that therapeutics designed with the protein encoded for by this gene could also be beneficial in the treatment of inflammatory skin diseases such as psoriasis, contact dermatitis, skin infection. Finally, high NOV1 gene expression is seen in the thymus suggesting that this gene product might be involved in the normal homeostasis of the thymus.

References:

1. Stepien H, Agro A, Crossley J, Padol I, Richards C, Stanisz A. Immunomodulatory properties of diazepam-binding inhibitor: effect on human interleukin-6 secretion, lymphocyte proliferation and natural killer cell activity in vitro. Neuropeptides 1993 Sep;25(3):207-11

NOV2

Expression of NOV2 was assessed using the primer-probe sets Ag1865, Ag2813, Ag1357, Ag1377, Ag2029, Ag2039, Ag2864 and Ag2747, described in Tables 5A, 5B, 5C, 5D, 5E, 5F, 5G and 5H. Results of the RTQ-PCR runs are shown in Tables 5I, 5J, 5K, 5L and 5M.

Table 5A. Probe Name Ag1865

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-AGGCAAAATCATCAACATCAAC-3'	22	1464	39
Probe	TET-5'-CTCAGCCCACCTCACAGAGACCATCT-3'-TAMRA	26	1500	40
Reverse	5'-TTAGCACACCAGGAGACATCTC-3'	22	1538	41

Table 5B. Probe Name Ag2813

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-AATCATCAACATCAACATTGCA-3'	22	1470	42
Probe	TET-5'-CTCAGCCCACCTCACAGAGACCATCT-3'-TAMRA	26	1500	43
Reverse	5'-GTTAGCACACCAGGAGACATCT-3'	22	1539	44

Table 5C. Probe Name Ag1357

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-ATCAGAACTCCTGCCACTCTTT-3'	22	487	45
Probe	TET-5'-TGGACCTATGACACGCAGCAATTCTT-3'-TAMRA	26	438	46
Reverse	5'-ATGCCAATGACTGAGAAAGTTG-3'	22	416	47

Table 5D. Probe Name Ag1377

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-TATTACTTGGGTGGTCATTCCA-3'	22	1037	48
Probe	TET-5'-CAACCCATGGAAAATTCTGGATTTCG-3'-TAMRA	26	1061	49
Reverse	5'-ATATTCCCAATGTTGCCATTTC-3'	22	1110	50

Table 5E. Probe Name Ag2029

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-AGGCAAAATCATCAACATCAAC-3'	22	1464	51
Probe	TET-5'-CTCAGCCCACCTCACAGAGACCATCT-3'-TAMRA	26	1500	52
Reverse	5'-TTAGCACACCAGGAGACATCTC-3'	22	1538	53

Table 5F. Probe Name Ag2039

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-GACCTATGACACGCAGCAAT-3'	20	442	54
Probe	TET-5'-TCTTCAACTTTCTCAGTCATTGGCAT-3'-TAMRA	26	416	55
Reverse	5'-GGAAGCCATGATTGCATATG-3'	20	367	56

Table 5G. Probe Name Ag2864

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-TGGCAGGAGTTCTGATATAACC-3'	22	493	57
Probe	TET-5'-TCAGTCCGACTGGAGAAAATCTCTAA-3'-TAMRA	26	515	58
Reverse	5'-GCGTTTGGAGTAGAAGTGAGAA-3'	22	564	59

Table 5H. Probe Name Ag2747

Primers	Sequences	Length	Start Position	SEQ ID NO:
Forward	5'-TGGCAGGAGTTCTGATATAACC-3'	22	493	60
Probe	TET-5'-TCAGTCCGACTGGAGAAAATCTCTAA-3'-TAMRA	26	515	61
Reverse	5'-GCGTTTGGAGTAGAAGTGAGAA-3'	22	564	62

Table 5I. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag1865, Run 207929096	Rel. Exp.(%) Ag2813, Run 209052394	Tissue Name	Rel. Exp.(%) Ag1865, Run 207929096	Rel. Exp.(%) Ag2813, Run 209052394
AD 1 Hippo	13.4	14.1	Control (Path) 3 Temporal Ctx	9.6	11.9
AD 2 Hippo	31.0	33.4	Control (Path) 4 Temporal Ctx	27.5	35.6
AD 3 Hippo	9.0	12.1	AD 1 Occipital Ctx	17.9	22.5
AD 4 Hippo	12.2	14.0	AD 2 Occipital Ctx (Missing)	0.0	0.0
AD 5 Hippo	90.1	**100.0**	AD 3 Occipital Ctx	10.9	11.6
AD 6 Hippo	55.9	58.6	AD 4 Occipital Ctx	22.1	25.5
Control 2 Hippo	30.8	33.7	AD 5 Occipital Ctx	47.0	49.3
Control 4 Hippo	12.2	19.9	AD 5 Occipital Ctx	32.1	36.6
Control (Path) 3 Hippo	11.3	11.3	Control 1 Occipital Ctx	7.4	6.3
AD 1 Temporal Ctx	21.3	27.4	Control 2 Occipital Ctx	46.0	48.6

AD 2 Temporal Ctx	32.3	33.2	Control 3 Occipital Ctx	15.7	19.5
AD 3 Temporal Ctx	6.3	10.2	Control 4 Occipital Ctx	13.6	12.8
AD 4 Temporal Ctx	24.1	24.3	Control (Path) 1 Occipital Ctx	96.6	79.6
AD 5 Inf Temporal Ctx	**100.0**	**100.0**	Control (Path) 2 Occipital Ctx	10.9	13.1
AD 5 Sup Temporal Ctx	53.6	55.5	Control (Path) 3 Occipital Ctx	8.5	6.5
AD 6 Inf Temporal Ctx	52.1	55.1	Control (Path) 4 Occipital Ctx	13.4	17.7
AD 6 Sup Temporal Ctx	48.3	53.6	Control 1 Parietal Ctx	12.0	12.6
Control 1 Temporal Ctx	7.7	11.7	Control 2 Parietal Ctx	46.0	57.8
Control 2 Temporal Ctx	31.4	37.9	Control 3 Parietal Ctx	14.2	18.6
Control 3 Temporal Ctx	18.0	20.4	Control (Path) 1 Parietal Ctx	58.2	71.2
Control 3 Temporal Ctx	10.4	12.4	Control (Path) 2 Parietal Ctx	25.9	29.1
Control (Path) 1 Temporal Ctx	51.1	51.1	Control (Path) 3 Parietal Ctx	12.8	11.0
Control (Path) 2 Temporal Ctx	28.7	38.7	Control (Path) 4 Parietal Ctx	33.9	45.7

Table 5J. Panel 1.2

Tissue Name	Rel. Exp.(%) Ag1357, Run 134665778	Rel. Exp.(%) Ag1377, Run 134831870	Tissue Name	Rel. Exp.(%) Ag1357, Run 134665778	Rel. Exp.(%) Ag1377, Run 134831870
Endothelial cells	2.6	3.7	Renal ca. 786-0	33.4	32.3

Heart (Fetal)	2.0	2.8	Renal ca. A498	8.6	8.7
Pancreas	2.8	2.6	Renal ca. RXF 393	2.4	1.6
Pancreatic ca. CAPAN 2	5.4	4.7	Renal ca. ACHN	4.8	3.1
Adrenal gland	15.5	16.7	Renal ca. UO-31	1.8	1.3
Thyroid	8.2	8.0	Renal ca. TK-10	4.2	2.4
Salivary gland	17.4	23.5	Liver	13.5	21.8
Pituitary gland	7.0	3.9	Liver (fetal)	3.4	5.3
Brain (fetal)	5.9	7.1	Liver ca. (hepatoblast) HepG2	16.5	12.2
Brain (whole)	22.5	19.3	Lung	4.0	5.0
Brain (amygdala)	10.5	12.3	Lung (fetal)	2.6	4.7
Brain (cerebellum)	8.7	10.1	Lung ca. (small cell) LX-1	20.7	16.7
Brain (hippocampus)	14.2	24.7	Lung ca. (small cell) NCI-H69	5.1	6.2
Brain (thalamus)	12.4	15.0	Lung ca. (s.cell var.) SHP-77	2.3	1.9
Cerebral Cortex	17.9	24.5	Lung ca. (large cell)NCI-H460	14.4	16.0
Spinal cord	9.0	7.4	Lung ca. (non-sm. cell) A549	9.2	9.3
glio/astro U87-MG	8.7	12.7	Lung ca. (non-s.cell) NCI-H23	3.5	1.9
glio/astro U-118-MG	3.1	2.2	Lung ca. (non-s.cell) HOP-62	4.3	2.9
astrocytoma SW1783	2.4	3.5	Lung ca. (non-s.cl) NCI-H522	13.6	11.8
neuro*; met SK-N-AS	47.6	**100.0**	Lung ca. (squam.) SW 900	8.8	8.0
astrocytoma SF-539	4.7	2.9	Lung ca. (squam.) NCI-H596	1.0	1.6
astrocytoma SNB-75	1.2	0.8	Mammary gland	4.8	4.4
glioma SNB-19	13.2	5.7	Breast ca.* (pl.ef) MCF-7	44.1	39.8
glioma U251	0.9	0.7	Breast ca.* (pl.ef) MDA-	7.8	14.6

			MB-231		
glioma SF-295	1.7	0.9	Breast ca.* (pl. ef) T47D	11.7	22.4
Heart	38.2	48.3	Breast ca. BT-549	12.3	15.8
Skeletal muscle	23.0	43.2	Breast ca. MDA-N	32.8	23.0
Bone marrow	2.1	3.3	Ovary	1.1	0.9
Thymus	2.9	2.6	Ovarian ca. OVCAR-3	7.2	4.3
Spleen	3.8	4.9	Ovarian ca. OVCAR-4	7.4	4.7
Lymph node	2.5	3.7	Ovarian ca. OVCAR-5	18.9	13.4
Colorectal	0.8	1.6	Ovarian ca. OVCAR-8	9.3	20.0
Stomach	7.4	6.3	Ovarian ca. IGROV-1	14.3	6.9
Small intestine	16.6	22.2	Ovarian ca. (ascites) SK-OV-3	8.7	5.6
Colon ca. SW480	0.8	0.7	Uterus	7.2	6.9
Colon ca.* SW620 (SW480 met)	4.2	8.5	Placenta	14.7	30.1
Colon ca. HT29	14.1	14.4	Prostate	6.2	10.3
Colon ca. HCT-116	17.0	6.6	Prostate ca.* (bone met) PC-3	12.2	14.3
Colon ca. CaCo-2	21.5	21.6	Testis	5.9	4.1
CC Well to Mod Diff (ODO3866)	3.2	3.2	Melanoma Hs688(A).T	1.6	1.5
Colon ca. HCC-2998	**100.0**	42.9	Melanoma* (met) Hs688(B).T	0.6	1.0
Gastric ca. (liver met) NCI-N87	25.0	31.2	Melanoma UACC-62	1.8	1.4
Bladder	7.6	9.3	Melanoma M14	4.1	2.1
Trachea	3.7	4.0	Melanoma LOX IMVI	0.9	0.8
Kidney	18.7	25.2	Melanoma* (met) SK-MEL-5	4.9	5.6

Kidney (fetal)	8.4	9.3			

Table 5K. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag1865, Run 148610768	Rel. Exp.(%) Ag2029, Run 165552311	Rel. Exp.(%) Ag2039, Run 146127643	Rel. Exp.(%) Ag2813, Run 165528059	Rel. Exp.(%) Ag2864, Run 165517732
Liver adenocarcinoma	10.2	4.3	13.5	9.1	0.0
Pancreas	5.0	8.3	6.0	7.0	2.1
Pancreatic ca. CAPAN 2	8.9	20.3	21.5	24.8	10.2
Adrenal gland	13.2	17.8	13.2	17.8	1.0
Thyroid	9.7	15.1	14.0	11.2	2.5
Salivary gland	3.9	8.1	6.0	11.3	2.1
Pituitary gland	7.6	9.6	10.9	8.8	4.2
Brain (fetal)	10.2	34.9	12.4	36.3	41.2
Brain (whole)	43.5	**100.0**	40.9	**100.0**	**100.0**
Brain (amygdala)	36.1	61.1	22.1	44.8	31.4
Brain (cerebellum)	10.0	54.3	10.4	38.2	80.1
Brain (hippocampus)	**100.0**	57.4	52.9	63.7	38.4
Brain (substantia nigra)	12.4	43.2	13.0	45.4	35.4
Brain (thalamus)	34.4	72.2	28.9	81.8	80.7
Cerebral Cortex	50.0	13.6	63.7	11.6	15.8
Spinal cord	19.6	68.8	35.1	44.1	11.8
glio/astro U87-MG	6.7	13.5	27.5	10.7	10.4
glio/astro U-118-MG	13.5	32.8	26.8	19.9	17.1
astrocytoma SW1783	8.7	13.2	25.2	11.2	8.1
neuro*; met SK-N-AS	52.9	56.6	**100.0**	51.8	41.2
astrocytoma SF-539	9.6	11.8	21.3	14.5	17.4
astrocytoma SNB-75	21.8	22.8	54.3	20.3	3.3
glioma SNB-19	11.2	17.1	24.5	22.7	7.0
glioma U251	3.2	19.8	6.8	14.3	10.3
glioma SF-295	10.3	6.6	9.6	8.7	3.4
Heart (Fetal)	6.4	1.0	9.0	0.7	0.0
Heart	7.5	30.4	12.9	34.4	2.9
Skeletal muscle	19.9	0.0	25.5	0.4	2.7

(Fetal)					
Skeletal muscle	6.7	73.7	11.1	52.5	4.2
Bone marrow	4.0	6.3	5.1	5.4	2.1
Thymus	4.5	5.1	8.4	6.4	2.4
Spleen	14.7	12.6	14.1	11.6	6.9
Lymph node	9.2	23.3	9.2	21.8	9.7
Colorectal	8.4	3.1	7.4	3.9	0.0
Stomach	20.0	17.3	15.4	11.4	3.0
Small intestine	12.8	31.4	12.5	18.9	16.4
Colon ca. SW480	10.0	3.9	17.3	4.3	0.0
Colon ca.* SW620 (SW480 met)	8.3	7.7	14.2	9.7	14.9
Colon ca. HT29	12.3	6.6	22.8	9.6	2.6
Colon ca. HCT-116	6.8	6.7	17.9	9.0	6.0
Colon ca. CaCo-2	25.3	16.4	49.7	17.4	0.0
CC Well to Mod Diff (ODO3866)	8.1	16.0	0.0	14.6	6.0
Colon ca. HCC-2998	20.9	18.9	43.5	19.6	11.7
Gastric ca. (liver met) NCI-N87	20.0	31.2	68.8	30.8	10.0
Bladder	11.3	17.7	13.9	11.8	7.8
Trachea	13.2	10.9	18.0	11.0	3.3
Kidney	7.9	23.8	7.5	18.2	6.3
Kidney (fetal)	6.7	13.5	9.5	14.8	5.2
Renal ca. 786-0	24.0	30.4	44.8	28.9	33.0
Renal ca. A498	24.5	37.9	48.0	36.9	36.3
Renal ca. RXF 393	5.2	15.4	6.9	19.6	21.9
Renal ca. ACHN	10.6	13.7	26.8	12.7	7.7
Renal ca. UO-31	12.3	20.2	15.3	16.8	17.2
Renal ca. TK-10	3.2	12.1	10.4	8.8	3.4
Liver	10.7	23.0	9.9	30.1	5.2
Liver (fetal)	6.9	19.3	6.5	16.4	3.1
Liver ca. (hepatoblast) HepG2	13.0	34.2	39.0	33.0	6.1
Lung	9.0	12.9	11.3	11.9	6.1
Lung (fetal)	12.6	13.6	18.4	19.6	0.0
Lung ca. (small cell) LX-1	14.6	29.9	33.4	29.7	9.2
Lung ca. (small	10.4	6.0	19.1	7.8	2.2

cell) NCI-H69					
Lung ca. (s.cell var.) SHP-77	11.5	23.8	31.4	15.9	29.9
Lung ca. (large cell)NCI-H460	5.2	25.2	8.1	21.8	17.0
Lung ca. (non-sm. cell) A549	12.9	8.0	23.5	14.6	7.1
Lung ca. (non-s.cell) NCI-H23	14.9	8.9	21.3	5.3	4.3
Lung ca. (non-s.cell) HOP-62	4.3	5.1	6.9	8.2	5.4
Lung ca. (non-s.cl) NCI-H522	24.1	11.0	23.2	14.8	10.2
Lung ca. (squam.) SW 900	9.0	19.9	28.1	20.7	7.5
Lung ca. (squam.) NCI-H596	3.0	6.3	6.6	6.9	5.6
Mammary gland	11.2	14.2	24.0	9.6	6.8
Breast ca.* (pl.ef) MCF-7	13.5	34.6	44.4	21.9	6.3
Breast ca.* (pl.ef) MDA-MB-231	15.0	46.7	51.8	38.2	10.5
Breast ca.* (pl. ef) T47D	7.0	8.7	14.3	6.7	6.8
Breast ca. BT-549	18.0	34.9	38.2	28.3	20.3
Breast ca. MDA-N	13.1	8.3	49.0	11.4	4.7
Ovary	9.0	0.6	6.7	0.8	0.0
Ovarian ca. OVCAR-3	6.3	17.2	22.4	20.9	12.7
Ovarian ca. OVCAR-4	2.8	17.6	4.7	13.7	8.7
Ovarian ca. OVCAR-5	20.3	23.2	27.0	21.3	2.4
Ovarian ca. OVCAR-8	11.7	7.4	17.1	9.1	6.7
Ovarian ca. IGROV-1	7.6	6.7	14.3	7.8	7.2
Ovarian ca. (ascites) SK-OV-3	15.8	24.1	0.0	37.4	12.1
Uterus	5.6	20.9	7.7	17.4	1.8
Placenta	6.2	14.9	12.7	8.4	3.7
Prostate	7.4	13.3	8.8	5.8	0.6
Prostate ca.*	8.2	7.9	13.5	9.9	3.6

(bone met) PC-3					
Testis	58.2	18.9	38.4	36.3	0.7
Melanoma Hs688(A).T	7.4	4.7	35.4	4.2	2.3
Melanoma* (met) Hs688(B).T	8.1	4.2	36.3	4.0	2.3
Melanoma UACC-62	0.4	2.5	1.1	2.9	3.9
Melanoma M14	9.5	32.3	10.7	34.6	6.7
Melanoma LOX IMVI	2.3	1.4	4.3	2.0	0.0
Melanoma* (met) SK-MEL-5	7.1	8.8	15.1	12.2	7.5
Adipose	8.6	9.3	10.7	8.0	9.0

Table 5L. Panel 2D

Tissue Name	Rel. Exp.(%) Ag2813, Run 162599032	Tissue Name	Rel. Exp.(%) Ag2813, Run 162599032
Normal Colon	56.6	Kidney Margin 8120608	4.5
CC Well to Mod Diff (ODO3866)	18.8	Kidney Cancer 8120613	11.2
CC Margin (ODO3866)	14.4	Kidney Margin 8120614	6.8
CC Gr.2 rectosigmoid (ODO3868)	13.4	Kidney Cancer 9010320	6.8
CC Margin (ODO3868)	3.4	Kidney Margin 9010321	11.9
CC Mod Diff (ODO3920)	10.9	Normal Uterus	3.5
CC Margin (ODO3920)	9.5	Uterine Cancer 064011	18.8
CC Gr.2 ascend colon (ODO3921)	28.1	Normal Thyroid	21.3
CC Margin (ODO3921)	9.8	Thyroid Cancer	16.8
CC from Partial Hepatectomy (ODO4309) Mets	60.7	Thyroid Cancer A302152	20.3
Liver Margin (ODO4309)	56.6	Thyroid Margin A302153	21.8
Colon mets to lung (OD04451-01)	8.7	Normal Breast	16.6
Lung Margin (OD04451-02)	8.1	Breast Cancer	13.1
Normal Prostate 6546-1	**100.0**	Breast Cancer (OD04590-01)	51.1

Prostate Cancer (OD04410)	31.6	Breast Cancer Mets (OD04590-03)	54.3
Prostate Margin (OD04410)	27.0	Breast Cancer Metastasis	37.9
Prostate Cancer (OD04720-01)	10.0	Breast Cancer	11.3
Prostate Margin (OD04720-02)	19.5	Breast Cancer	13.7
Normal Lung	35.6	Breast Cancer 9100266	24.0
Lung Met to Muscle (ODO4286)	20.4	Breast Margin 9100265	6.8
Muscle Margin (ODO4286)	8.6	Breast Cancer A209073	17.0
Lung Malignant Cancer (OD03126)	29.7	Breast Margin A2090734	8.4
Lung Margin (OD03126)	26.6	Normal Liver	36.9
Lung Cancer (OD04404)	17.9	Liver Cancer	19.2
Lung Margin (OD04404)	10.9	Liver Cancer 1025	15.7
Lung Cancer (OD04565)	4.7	Liver Cancer 1026	4.8
Lung Margin (OD04565)	5.0	Liver Cancer 6004-T	18.0
Lung Cancer (OD04237-01)	32.1	Liver Tissue 6004-N	9.1
Lung Margin (OD04237-02)	19.5	Liver Cancer 6005-T	7.6
Ocular Mel Met to Liver (ODO4310)	9.8	Liver Tissue 6005-N	4.9
Liver Margin (ODO4310)	44.8	Normal Bladder	25.9
Melanoma Metastasis	9.4	Bladder Cancer	4.6
Lung Margin (OD04321)	18.6	Bladder Cancer	11.1
Normal Kidney	69.3	Bladder Cancer (OD04718-01)	31.4
Kidney Ca, Nuclear grade 2 (OD04338)	58.6	Bladder Normal Adjacent (OD04718-03)	9.5
Kidney Margin (OD04338)	24.7	Normal Ovary	1.7
Kidney Ca Nuclear grade 1/2 (OD04339)	17.3	Ovarian Cancer	18.7
Kidney Margin (OD04339)	58.2	Ovarian Cancer (OD04768-07)	65.5
Kidney Ca, Clear cell type (OD04340)	36.6	Ovary Margin (OD04768-08)	4.4
Kidney Margin (OD04340)	33.9	Normal Stomach	12.9

Kidney Ca, Nuclear grade 3 (OD04348)	9.8	Gastric Cancer 9060358	3.4
Kidney Margin (OD04348)	29.3	Stomach Margin 9060359	12.2
Kidney Cancer (OD04622-01)	19.1	Gastric Cancer 9060395	13.8
Kidney Margin (OD04622-03)	4.6	Stomach Margin 9060394	9.8
Kidney Cancer (OD04450-01)	49.7	Gastric Cancer 9060397	36.1
Kidney Margin (OD04450-03)	44.4	Stomach Margin 9060396	2.6
Kidney Cancer 8120607	3.1	Gastric Cancer 064005	70.7

Table 5M. Panel 4D

Tissue Name	Rel. Exp.(%) Ag1377, Run 145703126	Rel. Exp.(%) Ag1865, Run 164038695	Rel. Exp.(%) Ag2029, Run 161379206	Rel. Exp.(%) Ag2039, Run 156265871	Rel. Exp.(%) Ag2747, Run 162015214	Rel. Exp.(%) Ag2813, Run 162293678
Secondary Th1 act	20.2	22.4	15.8	24.7	11.6	17.2
Secondary Th2 act	14.7	25.7	17.6	18.7	13.2	26.8
Secondary Tr1 act	21.2	15.5	20.7	21.9	9.7	18.9
Secondary Th1 rest	7.9	10.2	4.5	9.2	4.5	8.4
Secondary Th2 rest	12.6	8.6	9.3	7.1	4.6	10.8
Secondary Tr1 rest	8.0	14.8	9.8	10.2	4.6	13.8
Primary Th1 act	22.8	20.3	16.0	15.0	8.8	18.4
Primary Th2 act	23.3	28.1	15.6	18.6	12.3	18.4
Primary Tr1 act	30.8	24.1	17.2	30.1	13.9	26.4
Primary Th1 rest	46.7	49.0	33.7	65.5	21.0	49.0
Primary Th2 rest	19.2	17.7	18.6	29.7	9.2	23.8
Primary Tr1 rest	14.1	18.2	13.8	20.3	8.8	22.2
CD45RA CD4 lymphocyte act	12.1	17.8	9.3	13.2	5.4	13.5
CD45RO CD4 lymphocyte act	20.7	33.9	14.3	21.3	9.3	24.0
CD8 lymphocyte act	16.6	17.2	13.3	12.2	5.1	18.2
Secondary CD8 lymphocyte rest	18.3	22.4	16.2	13.1	6.9	20.0
Secondary CD8 lymphocyte act	14.4	17.4	12.1	14.0	4.6	13.1
CD4 lymphocyte none	6.5	9.2	6.8	7.5	3.0	5.7

2ry Th1/Th2/Tr1_anti-CD95 CH11	19.5	17.1	13.0	16.4	5.5	20.2
LAK cells rest	35.1	33.7	24.8	27.9	9.5	37.1
LAK cells IL-2	30.6	34.6	22.7	20.6	14.6	30.6
LAK cells IL-2+IL-12	21.3	21.3	15.2	14.6	9.5	17.8
LAK cells IL-2+IFN gamma	31.4	36.9	23.7	30.1	21.2	34.4
LAK cells IL-2+ IL-18	20.9	21.8	17.9	33.2	10.9	26.2
LAK cells PMA/ionomycin	9.4	13.5	11.0	15.7	3.7	13.9
NK Cells IL-2 rest	11.7	32.5	17.7	22.5	20.7	28.9
Two Way MLR 3 day	19.9	56.3	21.9	24.0	14.8	26.1
Two Way MLR 5 day	12.4	18.7	14.3	14.4	8.3	17.9
Two Way MLR 7 day	8.6	13.4	6.9	10.5	4.3	10.6
PBMC rest	9.9	18.0	9.5	12.2	4.5	12.9
PBMC PWM	70.7	85.9	53.6	58.2	48.0	74.7
PBMC PHA-L	35.4	31.9	14.5	23.5	17.0	24.5
Ramos (B cell) none	27.5	21.6	16.8	17.9	17.6	23.3
Ramos (B cell) ionomycin	72.2	97.3	80.7	97.3	**100.0**	**100.0**
B lymphocytes PWM	57.8	**100.0**	**100.0**	**100.0**	55.1	97.9
B lymphocytes CD40L and IL-4	15.0	23.5	20.0	24.8	30.4	28.9
EOL-1 dbcAMP	7.0	18.2	12.1	12.2	3.4	15.2
EOL-1 dbcAMP PMA/ionomycin	9.0	19.1	14.9	14.1	5.3	18.2
Dendritic cells none	19.6	35.8	21.3	21.2	9.2	29.7
Dendritic cells LPS	15.5	36.3	36.3	39.5	7.5	43.8
Dendritic cells anti-CD40	13.9	39.5	24.3	25.5	7.5	21.6
Monocytes rest	21.2	28.7	19.9	19.1	7.2	20.3
Monocytes LPS	22.1	30.1	24.0	21.2	22.8	21.9
Macrophages rest	26.4	38.2	30.8	30.6	15.3	34.4
Macrophages LPS	61.1	44.4	29.5	41.5	14.0	31.4
HUVEC none	11.2	17.9	17.6	18.3	8.2	19.9
HUVEC starved	26.2	38.7	19.9	36.1	15.3	34.6

HUVEC IL-1beta	4.9	12.9	7.3	11.7	4.4	7.1
HUVEC IFN gamma	14.2	27.9	17.6	17.8	10.2	24.5
HUVEC TNF alpha + IFN gamma	6.5	28.5	21.0	18.3	13.7	27.0
HUVEC TNF alpha + IL4	8.8	19.2	14.7	63.7	8.6	14.9
HUVEC IL-11	7.1	11.1	7.4	8.8	3.4	7.6
Lung Microvascular EC none	13.6	32.1	21.5	15.0	8.0	26.8
Lung Microvascular EC TNFalpha + IL-1beta	11.0	28.3	20.9	17.4	4.8	24.8
Microvascular Dermal EC none	34.4	35.4	31.6	29.5	8.5	38.7
Microsvasular Dermal EC TNFalpha + IL-1beta	15.7	25.0	18.9	21.3	11.0	23.5
Bronchial epithelium TNFalpha + IL1beta	11.4	20.0	2.0	0.7	11.0	18.7
Small airway epithelium none	3.6	6.8	6.1	6.4	4.4	6.9
Small airway epithelium TNFalpha + IL-1beta	26.6	40.3	30.1	44.1	20.6	35.4
Coronery artery SMC rest	13.1	17.7	13.8	20.9	6.8	21.2
Coronery artery SMC TNFalpha + IL-1beta	5.6	13.4	10.7	8.7	4.7	10.0
Astrocytes rest	4.7	18.6	13.7	12.9	5.2	16.7
Astrocytes TNFalpha + IL-1beta	8.0	15.1	10.8	8.3	4.7	11.7
KU-812 (Basophil) rest	6.8	9.9	7.4	5.4	7.9	9.3
KU-812 (Basophil) PMA/ionomycin	20.6	31.4	21.0	24.0	21.0	26.1
CCD1106 (Keratinocytes)	6.8	17.3	8.8	12.7	7.6	17.0

none						
CCD1106 (Keratinocytes) TNFalpha + IL-1beta	3.6	7.9	1.2	1.9	2.5	7.0
Liver cirrhosis	8.7	8.0	4.8	4.9	2.4	8.0
Lupus kidney	15.1	4.8	4.0	3.3	2.0	5.0
NCI-H292 none	27.2	39.2	37.4	34.9	7.3	40.3
NCI-H292 IL-4	34.2	54.0	43.5	75.3	10.4	54.0
NCI-H292 IL-9	30.8	52.9	48.0	59.0	7.0	57.0
NCI-H292 IL-13	18.0	40.6	28.9	31.2	6.8	16.3
NCI-H292 IFN gamma	13.6	49.3	39.8	52.1	14.3	48.0
HPAEC none	10.9	35.8	11.3	9.3	4.8	10.2
HPAEC TNF alpha + IL-1 beta	10.7	20.0	11.4	19.3	10.0	16.7
Lung fibroblast none	14.6	14.6	12.1	9.7	4.4	12.2
Lung fibroblast TNF alpha + IL-1 beta	9.1	21.0	12.7	15.1	4.8	15.8
Lung fibroblast IL-4	11.6	24.0	19.9	15.5	8.1	23.5
Lung fibroblast IL-9	9.9	24.7	19.3	14.5	10.2	17.9
Lung fibroblast IL-13	21.5	14.2	9.7	11.0	5.4	12.9
Lung fibroblast IFN gamma	22.8	42.6	39.0	37.9	17.7	37.9
Dermal fibroblast CCD1070 rest	20.9	37.6	31.6	32.3	12.9	34.2
Dermal fibroblast CCD1070 TNF alpha	39.5	70.2	60.3	63.7	27.4	65.1
Dermal fibroblast CCD1070 IL-1 beta	11.2	26.2	18.8	20.2	7.0	18.6
Dermal fibroblast IFN gamma	9.5	23.5	19.8	19.9	10.6	18.6
Dermal fibroblast IL-4	14.5	24.8	16.4	21.5	8.3	23.0
IBD Colitis 2	3.3	1.6	0.5	0.8	1.4	1.8
IBD Crohn's	3.4	4.5	2.7	2.9	0.1	4.2
Colon	26.6	52.5	41.8	51.8	15.0	53.2
Lung	10.9	24.1	17.1	20.3	6.0	23.5

Thymus	**100.0**	80.7	57.4	72.7	12.9	76.3
Kidney	19.9	28.7	25.5	23.8	4.9	27.5

CNS_neurodegeneration_v1.0 Summary: Ag1865/Ag2813

Results from two experiments using different probe/primer sets are in good agreement. A NOV2 gene is moderately expressed in all samples in this panel, confirming the expression of this gene in the brain. Gene expression levels are similar in all brain regions tested (hippocampus, temporal cortex, parietal cortex, and occipital cortex) and show no apparent alteration in patients with Alzheimer's disease. Please see panel 1.3d for potential utility of this gene in the central nervous system.

Panel 1.2 Summary: Ag1357/Ag1377

A NOV2 gene is expressed at moderate to high levels in all of the tissues on this panel, with highest expression detected in a brain cancer cell line and a colon cancer cell line (CTs=23.5). In general, expression of this gene appears to be higher in cancer cell lines than in normal tissues; this pattern of expression holds true for breast and ovarian cancer cell lines as well. Thus, expression of this gene may be useful for identifying colon or brain cancer. Furthermore, therapeutic modulation of the expression or function of this gene, through the use of small molecule drugs, antibodies or protein therapeutics, might be of benefit for the treatment of cancer.

Among tissues derived from the central nervous system, a NOV2 gene is expressed at high levels in fetal brain, amygdala, cerebellum, hippocampus, thalamus, cerebral cortex and spinal cord (CTs=25-27). High expression throughout the brain indicates a potential role for this gene in normal brain function. A NOV2 gene encodes a protein with homology to endozepine. Endogenous benzodiazepine-like substances are thought to play a role in the development of hepatic encephalopathy. Gooday R, Hayes PC, Bzeizi K, O'Carroll RE. Benzodiazepine receptor antagonism improves reaction time in latent hepatic encephalopathy. Psychopharmacology (Berl) 1995 Jun;119(3):295-8. Ohkuma S, Katsura M, Tsujimura A. Alterations in cerebral diazepam binding inhibitor expression in drug dependence: a possible biochemical alteration common to drug dependence. Life Sci 2001 Feb 2;68(11):1215-22.

It has been suggested that benzodiazepine receptor antagonism may improve cognitive function, particularly speed of information processing, in patients with latent hepatic encephalopathy. Thus, a NOV2 gene product used as a protein therapeutic, or drugs that stimulate the protein's function, may have efficacy in the treatment of hepatic encephalopathy.

Increased expression of diazepam binding inhibitor (DBI), a endozepine peptide with anxiogenic action, in Alzheimer's disease, addiction and schizophrenia, indicates that drugs that inhibit NOV2 protein activity may also have utility in the treatment of these diseases. Edgar PF, Schonberger SJ, Dean B, Faull RL, Kydd R, Cooper GJ. A comparative proteome analysis of hippocampal tissue from schizophrenic and Alzheimer's disease individuals. Mol Psychiatry 1999 Mar;4(2):173-8. Ohkuma S, Katsura M, Tsujimura A. Alterations in cerebral diazepam binding inhibitor expression in drug dependence: a possible biochemical alteration common to drug dependence. Life Sci 2001 Feb 2;68(11):1215-22.

Among tissues with metabolic function, this gene is expressed at high levels in thyroid, adrenal gland, pituitary gland, heart, skeletal muscle, and at moderate levels in pancreas. Therefore, this gene product may be important for the pathogenesis and treatment of diseases in any or all of these tissues. In particular, a diazepam binding inhibitor has previously been identified as a potential autoantigen in autoimmune diabetes in a screen of a human pancreatic islet cDNA library with the sera from the diabetic patients autoantigens in autoimmune diabetes (ref. 4). In addition, treatment of rodents with the octadecaneuropeptide [diazepam-binding inhibitor (33-50)] has been shown to result in decreased food intake and weight loss. de Mateos-Verchere JG, Leprince J, Tonon MC, Vaudry H, Costentin J. The octadecaneuropeptide [diazepam-binding inhibitor (33-50)] exerts potent anorexigenic effects in rodents. Eur J Pharmacol 2001 Mar 2;414(2-3):225-31.

These observations suggest that therapeutic modulation of the diazepam-binding inhibitor-like protein encoded by a NOV2 gene may be useful in the treatment of diabetes and obesity.

Panel 1.3D Summary: Ag1865/2029/2039/2813/Ag2864

Results from multiple experiments with different probe and primer sets show highest expression predominantly in regions of the brain. Among tissues derived from the central nervous system, a NOV2 gene is expressed at high levels in fetal brain, amygdala, cerebellum, hippocampus, thalamus, cerebral cortex and spinal cord (CTs=28-31). Please see Panel 1.2 summary for discussion of potential utility based upon expression in the CNS.

Among tissues with metabolic function, a NOV2 gene is expressed at low to moderate levels in pancreas, adrenal gland, pituitary gland, thyroid, heart, liver and adipose. Please see Panel 1.2 summary for discussion of potential utility of this gene in metabolic diseases.

Across all of the panels, there appears to be consistent expression of a NOV2 gene in the neuroblastoma cell line. This suggests that the expression of this gene could be used to

distinguish this cell type from others. Moreover, therapeutic modulation of this gene, through the use of small molecule drugs, antibodies or protein therapeutics might be of use in the treatment of neuroblastoma.

Panel 2D Summary: Ag2813

A NOV2 gene is most highly expressed in a sample derived from normal prostate tissue (CT=25.6). In addition, substantial expression of this gene is detected in a number of tissue samples on this panel. Strikingly, NOV2 gene expression is higher in cancers of the ovary, breast and stomach when compared with their associated normal adjacent tissues. Thus, the expression of this gene could be used to distinguish ovarian, breast or stomach cancer tissue from normal tissue. Moreover, therapeutic modulation of NOV2 gene expression or activity, through the use of small molecule drugs, antibodies or protein therapeutics might be of benefit in the treatment of ovarian, breast or stomach cancer.

Panel 4D Summary: Ag1377/Ag1865/Ag2029/Ag2039/Ag2747/Ag2813

Multiple experiments using different probe/primer sets gave results that are in very good agreement. A NOV2 gene is moderately expressed in the majority of samples on this panel (CT values ranging from 27.5 to 32). However, this gene is expressed at a higher level in activated B cells (Ramos, B cells plus PWM and PBMC plus PWM). A NOV2 gene encodes a protein with homology to a membrane-associated diazepam binding inhibitor, which has been shown to have immunomodulatory activity. Stepien H, Agro A, Crossley J, Padol I, Richards C, Stanisz A. Immunomodulatory properties of diazepam-binding inhibitor: effect on human interleukin-6 secretion, lymphocyte proliferation and natural killer cell activity in vitro. Neuropeptides 1993 Sep;25(3):207-11 Therefore small molecules target or antibodies against NOV2 proteins may modulate B cell activity and be useful in the treatment of diseases associated with B cell activation, such as autoiommune diseases (including systemic lupus erythematosus and rheumatoid arthritis) and hyperglobulinemia. In addition, a NOV2 gene is quite abundantly expressed in dermal fibroblasts treated with TNF-alpha, suggesting that therapeutics designed with the protein encoded for by this gene could also be beneficial in the treatment of inflammatory skin diseases such as psoriasis, contact dermatitis, skin infection. Finally, high NOV2 gene expression is seen in the thymus suggesting that this gene product might be involved in the normal homeostasis of the thymus.

OTHER EMBODIMENTS

Although particular embodiments have been disclosed herein in detail, this has been done by way of example for purposes of illustration only, and is not intended to be limiting with

respect to the scope of the appended claims, which follow. In particular, it is contemplated by the inventors that various substitutions, alterations, and modifications may be made to the invention without departing from the spirit and scope of the invention as defined by the claims. The choice of nucleic acid starting material, clone of interest, or library type is believed to be a matter of routine for a person of ordinary skill in the art with knowledge of the embodiments described herein. Other aspects, advantages, and modifications considered to be within the scope of the following claims.